AD_____

Award Number: DAMD17-94-J-4391

TITLE: Isolation of a Breast Cancer Tumor Suppressor Gene from Chromosome 3P

PRINCIPAL INVESTIGATOR: Harry Drabkin, M.D.

CONTRACTING ORGANIZATION: University of Colorado Health Science Center
Denver, Colorado 80262

REPORT DATE: October 1999

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1999	Final (23 Sep 94 - 22 Sep 99)

4. TITLE AND SUBTITLE
Isolation of a Breast Cancer Tumor Suppressor Gene from Chromosome 3P

5. FUNDING NUMBERS
DAMD17-94-J-4391

6. AUTHOR(S)
Harry Drabkin, M.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Colorado Health Sciences Center
Denver, Colorado 80262

e-mail:

harry.drabkin@uchsc.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release
distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

Loss of tumor suppressor genes (TSGs) represent critical molecular events in the development and progression of breast cancer. Based on loss of heterozygosity (LOH) studies as well as direct cytogenetic studies of breast tumors, one or more TSGs likely resides on the short arm of chromosome 3 (3p) and appears to be involved in nearly 50% of breast cancers. Four distinct regions within 3p [p12, p14, p21 (proximal) and p21 (distal)] undergo recurrent deletions in human carcinomas and are the most likely sites for a breast cancer TSG. In this Final Progress Report, we describe the full spectrum of investigations we have conducted to identify tumor suppressor genes and to begin to understand their functions. Early in our studies, we demonstrated recurrent homozygous deletion or rearrangement in breast cancer cell lines involving 3p14. The critical region was cloned and sequenced which led to the identification of several putative exons. We determined that 3p14 is subject to a high degree of genomic instability which is ongoing in some cases. We evaluated other 3p regions for involvement in breast cancer, as originally proposed, including the 3p21.31 and 3p21.33 homozgyous deletion regions as well as the interval between. Our studies led to the discovery of a novel *patched*-related gene, TRC8, which resides in a chromosomal region with frequent amplification in breast tumors. Current evidence suggests this gene will be important in early embryogenesis and may identify a new pathway for cancer development

14. SUBJECT TERMS
Breast Cancer, Gene Isolation

15. NUMBER OF PAGES
73

16. PRICE CODE

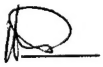
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

FOREWORD

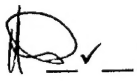
Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.


Where copyrighted material is quoted, permission has been obtained to use such material.

_____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

_____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

_____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).


__✓__ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.


__✓__ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.


__✓__ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.


In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.


10/21/99

PI - Signature Date

Table of Contents

5. INTRODUCTION

5.A. Purpose of the Work

Our project concerned the identification and isolation of a breast cancer tumor suppressor gene from the short arm of chromosome 3 (3p). Our Specific Aims were designed to: 1) define the regions of 3p undergoing LOH in breast cancer; 2) test known 3p candidate genes for mutations in breast tumors, 3) isolate additional candidate tumor suppressor genes from regions we defined and 4) characterize the product(s) of these genes and assess their involvement in breast cancer. We have made substantial progress towards these goals including the evaluation of several regions for involvement in breast cancers and the isolation of potentially important candidate genes. In the early stages of the investigation, we identified a region of homozygous deletion in a subset of breast cancer cell lines suggesting that a tumor suppressor gene would be found in the region of loss. The deletion occurred in a segment of DNA within a few hundred kilobases (1 kb = 1,000 base pairs, a measure of distance along the DNA molecule) of a chromosomal rearrangement involving chromosomes 3 and 8 which is associated with hereditary kidney carcinoma (Cohen et al., 1979). (It is often the case that a tumor suppressor gene is involved in more than one type of cancer, and both kidney and breast cancers are of epithelial cell origin.) The region surrounding this breakpoint (located in band 3p14.2) was one of the possible target loci described in our original application. We rearranged the original order of this project's specific aims to permit a detailed examination of 3p14.2, to develop necessary DNA reagents for thorough study of the target area and to identify potential genes. This focus allowed assembly of completed sequence for over 150 kilobases of DNA and identification of a number of potential gene coding regions (exons). Subsequently, we discovered that the deletion region in 3p14.2 was subject to onging genetic instability coincident with the common fragile site (FRA3B). We explored this instability phenomenon in some detail since it had the potential to provide important insights into mechanisms of genetic change which are a major feature of breast tumor cells. We found that ongoing instability was variable, as a breast tumor cell line had little instability while a cervical tumor line had a great deal. We also directed efforts at three additional deletion sites which could harbor important breast cancer related genes. We described a new gene in proximal 3p21 and developed a set of contiguous PAC clones covering the most distal homozygous deletion region in 3p21. (Distal and proximal refer to directions along the chromosome, with distal meaning towards the telomere, the structure which caps the chromosome arm and proximal meaning towards the centromere.) Our Final Progress Report reviews all these findings and places them into context with regard to our understanding of cancer.

5.B Nature of the Problem

Basic concepts especially for the lay reader We realize that the scientific literature is nearly totally composed of technical terms. The balance of the Introduction is designed to explain some of the major concepts guiding this project in terms understandable by lay individuals. Throughout the balance of this report, we have attempted to explain new concepts in similar lay terms.

The malignant potential of any tumor, including breast cancer, is a consequence of specific alterations (mutations, deletions, amplifications, etc.) in target genes that regulate the growth and biologic behavior of cells. (Genes are segments of DNA which encode proteins; DNA is "transcribed" into RNA and RNA in turn is "translated" into protein.) Whether cells grow slowly and remain localized, or proliferate rapidly and spread to distant sites (metastasize) is a complex process involving a host of regulatory genes. For example, loss or mutation of the p53 tumor suppressor gene, located on 17p, is associated with instability of the genome (entire DNA of the cell) (Livingstone et al., 1992) and a worsened prognosis. This instability results in an enhanced capacity of the malignant cell to undergo DNA rearrangements leading to alterations in critical regulatory genes. Loss of normal p53 function is also associated with the cell's ability to escape death or cell cycle arrest resulting from therapeutic radiation or chemotherapy (Lowe et al., 1993).

In other instances, the critical regulatory genes have yet to be identified. This is the case for genes located on 3p although some important candidates have been discovered in target regions. Scientific investigations have provided strong evidence pointing to where certain types of critical genes are likely to be located. For example, cytogenetic studies, which examine the content and nature of chromosomes within cells, have identified certain recurrent abnormalities in cancers. Specific chromosomal segments have been

found to be increased in number (amplified). This finding is expected to be associated with overexpression of a gene (because of its increased copy number). Such genes, for example, may encode growth factor receptors or may encode proteins that mediate resistance to chemotherapeutic agents. An example is provided by the MDM2 gene whose protein product inhibits the activity of p53; overexpression of MDM2 is thought to have consequences similar to mutation in p53. In contrast, cytogenetic studies have also pointed to recurrent deletions involving specific chromosomal regions. The critical genes believed to be encoded in these regions are referred to as tumor suppressor genes, the type of gene located on 3p which has been the focus of our investigation.

The nature of known tumor suppressor genes is quite varied. Certain tumor suppressor genes, e.g. p16 (an inhibitor of the cyclin dependent kinases or CDKs) and the retinoblastoma gene (RB1, an inhibitor of the E2F transcription factor), control cell division by regulating the process of DNA replication. Some proposed tumor suppressor genes, such as the chromosome 18 gene DCC (Deleted in Colon Carcinoma), are located on the cell surface. DCC has been shown to encode a receptor for a netrin, a protein involved in nerve growth cone development (Keino-Masu et al., 1996). Although a mouse mutation which destroys netrin function does not generate murine cancers, this may result from differences between mice and humans (Fazeli et al., 1997). One of the genes we identified in the proximal 3p21.31 deletion region is Semaphorin IV (now called SEMA3F) which is also involved in nerve growth cone guidance. This suggests that molecules initially identified in signaling pathways associated with nerve growth cone guidance may be involved in cancer. Importantly, Yamada et al., (Yamada et al., 1997) have recently demonstrated that a related semaphorin (semaphorin E) can cause non-MDR drug resistance in expressing cells. In addition, SEMA3F has been implicated (Naylor et al., 1998) as the gene responsible for the 3p21 driven tumor suppression observed in murine A9 cells (Killary et al., 1992). Of potentially major significance, SEMA3F is now known to interact with a receptor which also binds VEGF, an important regulator of new blood vessel formation. Thus SEMA3F could interfere with the generation of new blood supplies for growing tumors by inhibiting VEGF signaling. In this context, SEMA3F could act as a tumor suppressor gene. A primary feature of tumor suppressor genes is that their normal function is lost as part of tumor development. As a consequence, a regulatory function generally affecting growth and differentiation is also lost. The resulting cell may divide more frequently than is appropriate, giving rise to a clone or small cluster of related but abnormal cells and setting the stage for further genetic changes.

5.C. Methods to Isolate/Identify Tumor Suppressor Genes

5.C.1 Positional Cloning.

Tumor suppressor genes have been isolated by two approaches. Perhaps the most frequently used method is referred to as "positional cloning" in which the region of chromosomal loss is defined by molecular (DNA) probes and cytogenetic analysis. Because a visible chromosome deletion represents a large expanse of actual DNA, it is necessary to narrow the target region as much as possible. In one approach, this is done by using "polymorphic" probes which can distinguish between the two copies of the chromosome in question (each being inherited from one parent). In the tumor DNA, loss of one chromosomal copy is referred to as "loss of heterozygosity" (abbreviated LOH). To detect LOH, a DNA probe must exist which corresponds to the target DNA in question and, importantly, this bit of DNA must exhibit a frequent, naturally occurring, variation in the population. Such variations can usually be detected experimentally and they provide a means to identify differences between the two chromosome copies in any individual tumor sample. Naturally occurring differences in DNA sequence are not uncommon although some types of DNA sequences, referred to as "microsatellites", exhibit much more variation than other types. Microsatellites consist of repeated pairs of nucleotides (usually cytosine followed by adenosine, abbreviated CA) at specific chromosomal sites with the important feature that the number of pairs is variable from person to person. Since the natural variation of markers is often a limiting factor even with microsatellite markers, this explains the general necessity to test fairly large numbers of tumor samples with different DNA probes. The goals of LOH experiments are to identify one or more target regions and to narrow such regions as much as possible prior to performing gene searches.

5.C.2 Candidate Genes.

A second approach to identifying tumor suppressor genes is through the testing of candidate genes. These candidates come from two sources; first, genes mapped within regions of recurrent deletion are all considered candidates until proven otherwise, and second, those candidates whose known function suggests they might be targets independent of chromosomal position. A number of potentially important candidate loci have been identified in 3p target regions and we have tested some of these. These genes include the putative coding regions within 3p14.2, the SEMA3F and DEF-3 genes in 3p21.31 and cDNAs identified in 3p21.33. In addition, testing is ongoing for the *patched*-related gene TRC8, which is altered by a chromosomal translocation involving 3p14.2 and maps to a chromosomal region frequently amplified in breast tumors.

5.C.3 Identification of Homozygous Deletions Facilitates Positional Cloning of Tumor Suppressor Genes

LOH studies can lead to the identification of homozygous deletions in tumors, an extremely important finding which can greatly facilitate the precise positioning of tumor suppressor genes. Furthermore, homozygous deletions provide one of the best ways to chose particular candidate genes from a larger set, based upon their location within the minimal deletion defined experimentally. These often rare events thus provide a powerful adjunct to LOH and candidate gene studies.

A homozygous deletion means that both copies of a chromosome have undergone loss for a particular segment of DNA (see diagram below).

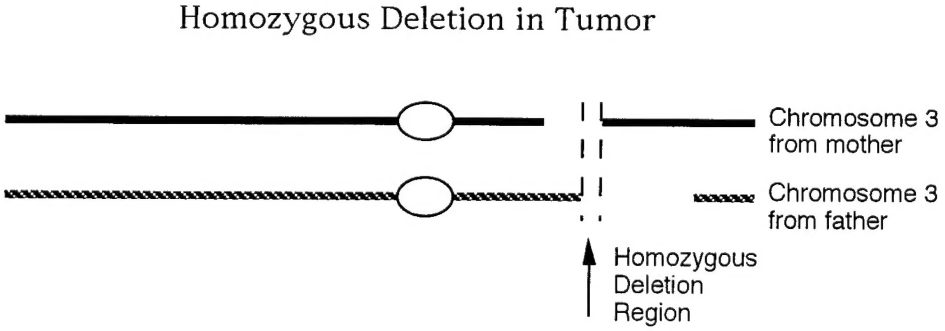


Figure 1. Diagram of chromosome 3 homologues with overlapping deletions in a tumor.

Such deletions represent one of several mechanisms that can lead to the complete loss of a tumor suppressor gene. While tumors frequently lose one tumor suppressor gene copy by undergoing large chromosomal deletions, the remaining copy is usually rendered non-functional by other genetic mechanisms. These can include point mutations and gene silencing (caused by hypermethylation) as well as a second deletion leaving a complete gap in the genetic material.

5.C.4 Chromosome 3 deletions in breast cancer.

In this section, we present a summary of data from the literature regarding genetic deletions of chromosome 3 in breast cancer. Sato et al. (Sato et al., 1991) examined 120 breast cancers for loss of heterozygosity (LOH) using a series of 3p polymorphic loci. Nearly 50% of the informative tumors (56/120) demonstrated LOH involving 3p. In this study, the 3p region undergoing the greatest loss was 3p14.2-p13 which includes the homozygous deletion region we identified. This study was one of the largest in terms of the number of tumors examined and 3p-derived probes tested. One significant limitation was that the type of probe used did not detect sufficient polymorphisms to allow the limits of the 3p target region to be defined more precisely. A cytogenetic study (chromosome analysis by direct microscopic visualization) by Pandis et al. (Pandis et al., 1993) identified 3p deletions in 5 of 41 breast carcinomas. These deletions appeared nearly identical and involved the 3p13(p14) region implicated above using LOH analysis. Intriguingly, in 3 of the 5 cases studied by Pandis et al., (Pandis et al., 1993) the 3p deletion was seen as the only recognizable cytogenetic abnormality. This suggests that 3p deletion may be an early event in a

subgroup of breast cancers. Another study using LOH analysis demonstrated that 3p loss was most frequently seen in familial breast cancers (Bergthorsson et al., 1995). The highest frequency of loss (68%) was with a probe (D3S1217) also located in the 3p14 region. These observations are significant since both BRCA1 and BRCA2 appear to be involved in recombination and DNA repair (Scully et al., 1997b; Scully et al., 1997a). Our previous results have suggested that 3p14 deletions result from the loss of an inherently unstable segment of DNA which coincides with the 3p14 common fragile site (Boldog et al., 1997). These deletions may be caused by recombinational errors, which would render this segment particularly sensitive to loss in a background of BRCA1 or 2 mutations. Taken together, there is considerable evidence from multiple investigators using different techniques which supports the frequent alteration of 3p14 during breast cancer development.

Additional data from LOH and cytogenetics studies implicate other 3p regions that may harbor genes important in breast cancer. Chen et al (Chen et al., 1994) found LOH within 3p13-14, 3p21-22 and 3p24-26 along with one case of homozygous deletion in 3p13 likely corresponding to our FRA3B associated homozygous deletions. All three of these target sites have been previously implicated in other carcinomas, particularly of lung. Hainsworth et al., (Hainsworth et al., 1991) found frequent chromosomal breakpoints at 3p21 which can be indicative of gene disruption or deletion. In a study by Ali et al, (Ali et al., 1989), the shortest region of common loss was between the markers D3F15S2 and RAF1, spanning the chromosomal region 3p21-p25 and a similar study by Devilee et al (Devilee et al., 1989) found the p14 to p21 region commonly deleted. Several other studies corroborate these findings having observed interstitial 3p deletions (Teixeira et al., 1996) or other structural changes affecting this chromosome (Lu et al., 1993). Intriguingly, Buchagen et al., (Buchhagen et al., 1994) identified one apparent rearrangement and one homozygous deletion for the probe D3S2, known to map within proximal 3p21.1 (Gemmill et al., 1991) very close to the SEMA3F deletion region (Roche et al., 1996). These studies all implicate multiple target regions on 3p which appear to be very similar to regions identified from studies on carcinomas derived from other tissues. Based on these studies, our efforts focused on three homozygous deletion regions. We demonstrated that 3p14 is inherently unstable in tumor cells with p53 mutations. We believe that instability is the primary event affecting 3p14, although it has not been possible to ruled out loss of a tumor suppressor function as well. Regardless, it may be possible to use the status of 3p14 as a biomarker for diagnostic or prognostic tests. We have examined two candidate genes in proximal 3p21 and have completed a search for homozygous deletions over an estimated 750 kb of the distal deletion region in 3p21. Our discovery of the TRC8 gene has generated new possibilities for investigation, although these will have to be supported by other mechanisms since this current project has ended.

6. BODY

In this section, we summarize the progress made throughout the four years of this project and provide references to publications. Some of the more recent results are being prepared for publication. Our most recent efforts have included investigating the functions and potential roles for genes in the proximal 3p21 region. In addition, our discovery of the *patched*-related gene, TRC8 (Gemmill et al., 1998), has led to important investigations to understand its role in cancers. While TRC8 does not map to 3p, it is located in 8q24, a region of frequent amplification in breast tumors.

6.A. 3p14 alterations in breast carcinomas.

6.A.1 Examination of breast cancer cell line DNAs for homozygous deletions in 3p14.

Southern blot hybridizations with probes from 3p14 detected homozygous deletions or rearrangements in DNAs from breast cancer cell lines (Boldog et al., 1997). Of the 13 breast carcinoma cell lines tested, one contained a discontinuous deletion while two were apparently rearranged. We developed a ~300 kb cosmid/lambda (DNA clone) contig in the region of the homozygous deletions (Boldog et al., 1997). From sequence and hybridization data, we showed that the cloned region corresponds to the most inducible common fragile site in the genome, FRA3B (Glover et al., 1984). Fragile sites are DNA regions which are unstable. In a few cases of rare fragile sites, their nature has been elucidated at the DNA sequence level and appears to be due, at least in part, to an expanded triplet repeat which may interfere with normal DNA replication. Our results clearly demonstrated that FRA3B represents a region of DNA rather than a single

site and does not contain an obvious triplet repeat. Where we have accurately defined the boundaries for the carcinoma-associated deletions, one or both are contained within FRA3B, showing that these deletions overlap the fragile site.

During our studies, Ohta et al. (Ohta et al., 1996) identified the FHIT gene from 3p14, reporting frequent abnormalities in RT-PCR products from carcinoma cell lines derived from colorectal and lung tumors. (In RT-PCR experiments, RNA is isolated and reverse transcribed into cDNA. PCR primers are then used to greatly amplify this product using the polymerase chain reaction which can then be examined for its correct size or its DNA sequence can be determined.) These observations were subsequently extended to breast carcinomas (Ahmadian et al., 1997; Panagopoulos et al., 1996; Negrini et al., 1996). In our previous Progress Reports, we outlined multiple reasons for skepticism about these results. Our own studies have led us to conclude that 3p14 is inherently unstable, particularly in tumors with p53 mutations (Boldog et al., 1997). This instability appears to be the primary driving force for deletions affecting 3p14. All of the data regarding the FHIT gene can be readily understood as the result of a bystander effect in which FHIT is damaged as a consequence of deletions within an unstable region. Whether or not these deletions confer a selective advantage to tumor cells is to date an unanswered question. However, more recent results from Dr. J Roth have demonstrated that FHIT can act in a suppressive manner (Ji et al., 1999). Whether or not this recapitulates events *in vivo* awaits further analysis.

6.A.2 DNA Sequencing and Analysis

We undertook DNA sequencing studies as a means to identify genes and to understand the instability of the region (Boldog et al., 1997). DNA sequencing was completed and deposited in the public databases for approximately 180 kb and our analysis of the sequence features has been published for a contiguous stretch of 110 kb (Boldog et al., 1997). (The DNA sequence analysis involves computer algorithms to identify all repetitive DNA sequences, identities and similarities to known genes contained in various databases, predicted gene segments and various other structural/compositional features.) We found the region to be very high in A-T base pair content with frequent LINE and MER repetitive elements, and conversely low in Alu repetitive elements and confirmed genes. In contrast to the reported rare folate-sensitive fragile sites, which are associated with expanded CGG repeats (Pintado et al., 1995; Nancarrow et al., 1995), FRA3B did not contain an expanded triplet repeat or methylated CpG-island. Inoue et al., (Inoue et al., 1997) subsequently sequenced 207 kb of this region which partially overlapped our data and extended the sequence by 130 kb in the telomeric direction. These additional data show the region continues to be AT rich and still lacks a triplet repeat. Our sequence data identified a number of putative exonic sequences which we investigated; the results are described briefly below.

6.A.3 Additional genes in FRA3B and 3p14?

Sequence data for 3p14.2 revealed a number of potential gene encoding regions within FRA3B in addition to FHIT. Many exons were predicted by GRAIL2 and GeneMark exon prediction programs (Borodovsky and McIninch, 1993), but nearly all occurred within repetitive elements of the LINE and MER families. Some showed identity with an EST from the dbEST sequence database, but these were discarded as candidate genes because they appeared to be primed from a genomic poly-A tract and were co-linear with genomic DNA. (ESTs are short DNA sequences obtained from the ends of cDNA clones. While most cDNA clones represent parts of genes, there are some which are derived from incompletely spliced RNA and others which may represent DNA contamination in the library.) Of five highly predicted exons which remained, none showed similarity to known genes. Several exons with lower probability scores occured in regions which were identified by other experimental evidence. Thus, this portion of the deletion region has the potential to encode more genes than FHIT.

We previously described results on several putative exons including λ58, GB and HRCA1. Despite the evidence from RT-PCR experiments suggesting expression of these three putative genes, all were found to be negative on Northern blots. We currently believe that these putative exons were expressed as read-through transcripts. The large amount of sequence data available from our efforts and from Inoue et al., (Inoue et al., 1997) provides a much more in depth analysis of the region. At this time, there appears to be

no compelling evidence that any additional genes are present in this region. We also undertook a series of functional tests to identify potential tumor suppressor genes within YAC clones (yeast artificial chromsomes) containing human DNA from 3p14. We introduced a selectable marker (by retrofitting) into YAC 74B2g and transfected the intact YAC into murine A9 fibrosarcoma cells. Initial testing of the tumorigenic potential of A9 parental and 74B2g transfected cells in *nude* mice suggested that the 74B2g transfectant was highly suppressed in its ability to form tumors. These experiments have now been repeated and no differences were observed. These observations illustrate the difficulties in performing such analyses. It is likely that the original evidence for tumor suppression was due to clonal variation.

6.A.4 Ongoing instability of 3p14.

The lack of genes in 3p14.2, other than FHIT, coupled with the problematic evidence used to support a tumor suppressive function for this gene, led us to consider alternative explanations for the deletions within this region. A likely hypothesis was that the presence of FRA3B sequences made the region inherently unstable in carcinoma cell lines. This possibility was suggested by the often discontinuous deletions we observed in carcinomas, including the breast tumor MDA231. We reasoned that if the region was inherently unstable, then carcinoma cell lines which still retained much of the region might be progressively deleting more of FRA3B and these events could be measured. Our strategy was to subclone two lines with known deletions and to analyze their DNA for alterations within FRA3B using PCR and Southern blots (see Figure 2 from our 1998 Progress Report). Originally, we found that 11 of 25 CC19 subclones (44%) had new deletions, no two of which were identical. This analysis was expanded to a total of 33 sub-clones of which 16 contained new deletions (48%). These results could have been explained by ongoing instability or by the sub-cloning of variants which were already present in the population. To distinguish between these possibilities, we performed a second sub-cloning experiment using a previous sub-clone of CC19 which showed no evidence for new deletion (CC19 sub-clone-7). We reasoned that if instability was truly ongoing, then a sub-clone without new deletions should show evidence for instability after a period of growth followed by sub-cloning once again. We grew CC19 sub-clone-7 for 10 passages, re-subcloned it and analyzed 64 sub-clones finding two (3%) with definitive evidence for new deletions. This low rate suggested that deletions in the original sub-clones had accumulated over time, while the presence of new deletions in CC19 sub-clone-7 derived cells confirmed that deletions were ongoing. In contrast, the breast tumor cell line MDA231 showed no evidence for new deletions after a similar period of culture and subcloning. Thus we identified a major difference between cell lines, both of which contain p53 alterations and 3p14 deletions. In one case the deletions are ongoing at a high rate while in the second the deletion appears stable. The nature of this fundamental difference is at present unknown.

Efforts to extend these results by examining the effects of activated RAS and MAP kinase kinase (MKK) pathways on this deletion rate are ongoing. Once this analysis is completed, it will be published as it provides important insight into the mechanism of deletions in this region. Activated RAS has been shown to induce chromosome breaks and other aberrations in NIH 3T3 cells with p53 mutations. In collaboration with Dr. Natalie Ahn (Univ. of Colorado, Boulder) we have introduced a mutant RAS expression construct into CC19 cells. The RAS mutation replaces the histidine at condon 61 with a leucine, rendering the protein GTPase deficient and constitutively active. RAS H61L has been incorporated into the pZIP mammalian expression vector and is expressed under control of the CMV promoter. RAS H61L transfectants of CC19 and control parental cells have both been subcloned (43 parental and 38 RAS transfectants subclones, respectively). We are testing DNA from these cells for the presence of new 3p14 deletions and to determine if activated RAS alters deletion rates.

6.A.5 Summary and Significance of 3p14 Findings

The identification of a DNA segment which undergoes recurrent homozygous deletion and rearrangement, as we have observed for this region of 3p14, initially suggested that it encoded a tumor suppressor gene. However, at the present time our data do not support FHIT as the 3p14.2 tumor suppressor . We proposed an alternative hypothesis; that genomic instability *per se* may lead to the observed deletions. While there are putative exons, there is no further support that these represent functional genes. These sequences have no matches in databases and we have observed no evidence for expression by Northern blots. To the best of our knowledge, this region undergoes loss as a result of genomic instability

and not through selection against a tumor suppressor gene (Boldog et al., 1997). We were the first to propose this alternative hypothesis for the origin of 3p14 deletions. It is based upon our detailed sequence analysis, gene searching, deletion analysis and functional assays, much of which has been supported under the auspices of this grant. We find it very interesting however that there are marked differences in the rates of *de novo* deletions involving this region in different tumor cells. These differences are not explained by p53 mutations and suggest additional alterations which influence genomic instability in this region. Such mechanisms would have the capacity to effect the development and progression of breast cancer. Our efforts to understand the specific mechanisms leading to instability of FRA3B will continue, despite the end of this project. The results should provide significant insights into this overall process which is such a key event in development of breast cancers. Furthermore, this understanding could lead to improved tests for detecting such changes in the early stages of carcinogenesis.

6.B. The TRC8 Gene

6.B.1 The t(3;8) breakpoint interrupts a novel *Patched*-related gene, TRC8.

Our results above suggested that the only gene within the 3p14/FRA3B region was FHIT. However, this left unanswered the question of what the 3;8 translocation was doing to induce hereditary cancer. Our discovery of the TRC8 gene, related to the Sonic Hedgehog (SHH) receptor, *patched* (PTCH), has begun to provide answers to this question.

The hereditary 3;8 translocation associated with 3p14 provided one of the strongest reasons for believing that this band contained a tumor suppressor gene. It had been assumed by many investigators that the translocation interrupted a tumor suppressor gene which was also deleted in many sporadic carcinomas including breast tumors. Although FHIT appeared to meet these expectations when first discovered, a substantial body of evidence now exists refuting this. As a result, we explored the alternative possibility that this translocation involved a gene on chromosome 8. The recent observation by Geurts et al., (Geurts et al., 1997) that FHIT was fused to the HMGI(C) gene in a benign parotid adenoma suggested that a similar fusion might result from the (3;8) translocation. To further our understanding of the genetic events which can affect 3p14, we undertook RACE experiments to identify any such gene fusions. We discovered that FHIT is indeed fused to a novel chromosome 8 gene, which we named TRC8 (Gemmill et al., 1998), and that this gene showed significant similarity to *patched* (PTCH), the receptor for the *sonic hedgehog* (SHH) signaling molecule. This is of major significance for several reasons. Mutations in PTCH are responsible for the basal cell nevus syndrome, a hereditary form of skin cancer, and are frequently found in sporadic basal cell carcinomas as well as medulloblastomas. Thus TRC8 is related to a known tumor suppressor gene and may function in an analogous fashion. However it is important to understand that we do not yet know if TRC8 will act as a recessive tumor suppressor or as a dominant oncogene. This region of chromosome 8 has been demonstrated by Wigler et al., (Lucito et al., 1998) to harbor several independent amplification domains in breast tumors. Thus it is possible that TRC8 is affected by one of these amplification domains and we are exploring this possibility. We have established a collaboration with Dr. Fujio Kasumi at the Japanese Foundation Cancer Institute in Tokyo who has provided matched tumor/normal DNA samples from breast tumors. These samples are being analyzed for TRC8 copy number and where possible, for levels of TRC8 expression. Such data should help establish whether or not TRC8 is a target for amplification in breast cancers. SAGE data available from the NCBI also supports this idea as a breast adenocarcinoma cell line shows significant (10X) overexpression of TRC8.

6.B.2. A Drosophila Homolog for TRC8.

We have discovered a homolog of TRC8 in Drosophila (dTRC8) and have obtained some direct evidence for function, albeit preliminary. Using the technique of double stranded RNA inhibition, we have shown that early embryogenesis in Drosophila is disrupted when dTRC8 is inhibited. Many embryos die with cuticle defects consisting of holes in their ventral surface. Complementary experiments with transgenic flies carrying an inducible version of the dTRC8 gene have demonstrated high levels of lethality when the gene is overexpressed. The resulting embryos also show cuticle defects but the spectrum is different, consisting mostly of anal structures with more anterior elements absent. Identification of the Drosophila homolog has provided us with an important opportunity to define the function of TRC8 in a simpler context.

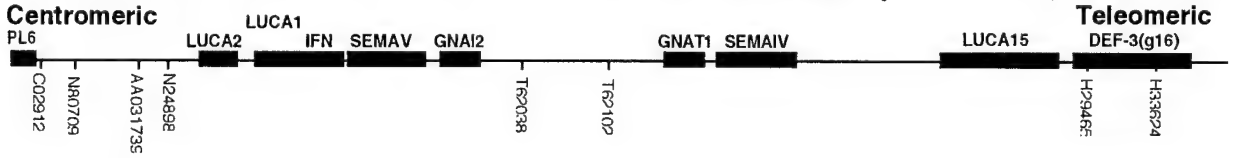
In particular, we are working to define the pathway that TRC8 belongs to and thus uncover an entire set of interacting genes which will need to be examined for alterations in carcinomas. Our discovery of TRC8 led to a funded grant proposal and to a recently submitted proposal on the Drosophila homolog.

6.C. Investigation of 3p21.31 in Breast Carcinomas.

6.C.1. Known and Predicted Genes in 3p21.31.

Significant evidence exists supporting the hypothesis that 3p21 harbors one or more potential tumor suppressor genes (reviewed in our 1997 Progress Report). First, the region has the highest frequency of LOH in many carcinomas and has been the target of homozygous deletions in cell lines and uncultured tumors. Importantly, the involvement of this region has been repeatedly suggested by LOH studies in breast cancer. Secondly, functional studies by Killary et al., (Killary et al., 1992) have shown that DNA segments from this region will suppress tumorigenicity of the mouse fibrosarcoma cell line, A9. Clone contigs and continuous DNA sequence data are available for this entire region and over 10 genes have been identified. The major genes in this interval are shown in the Figure below.

Two of these genes, SEMA3F (SEMAIV) and DEF-3, were isolated by our efforts (Roche et al., 1996;



Drabkin et al., 1999) and we have made substantial progress in understanding their structure and functions, as described later below.

6.C.2. Analysis by PCR for 3p21.31 Homozygous Deletions in Breast Carcinomas.

In our 1997 Progress Report, we described a detailed search for homozygous deletions in breast cancer lines using the polymerase chain reaction (PCR). PCR analyses were used to assess these lines for homozygous deletions in the 3p21.31 area corresponding to the map shown above. PCR results were obtained for 12 cell lines with 11 markers corresponding to each gene in the region. This analysis found no evidence of homozygous deletions with this set of cell lines and markers. However, another group has identified a homozygous deletion in a breast cancer cell line involving this region (Sekido et al., 1998).

6.C.3 Investigation of SEMA3F.

The rationale for studying this region derived from three overlapping homozygous deletions in small cell carcinomas, one of which we identified. Furthermore, most tumor suppressor genes are involved in more than one type of carcinoma, strongly suggesting that the same genes would be important in breast cancers. Within this deletion region, we as well as two other groups identified the SEMA3F gene (Roche et al., 1996; Sekido et al., 1996; Xiang et al., 1996) which encodes a member of the semaphorin/collapsin family. While the function of this gene family was initially associated with the migration of nerve growth cones, the broad expression patterns, gene knock-out studies and the finding that semaphorin E could mediate multidrug resistance (Yamada et al., 1997) argues that these genes are extremely important in non-neural development. Semaphorin E, related to SEMA3F, has recently been shown to be responsible for at least some drug -resistant phenotypes encountered in tumor cells (Yamada et al., 1997). Yamada et al., (Yamada et al., 1997) selected for cDNAs which were capable of conferring resistance to cis-platinum in Cos-7 cells. The primary clone recovered was semaphorin E and transfection of this clone could confer resistance to other cell lines.

The region of 3p21.3 including SEMA3F was shown by Killary et al., (Killary et al., 1992) to confer tumor suppression when introduced into murine A9 fibrosarcoma cells. In subsequent studies using individual P1 clones, SEMA3F was implicated but the data were inconsistent (Todd et al., 1996). More recent studies from the Naylor lab (Naylor et al., 1998) have provided additional evidence that SEMA3F

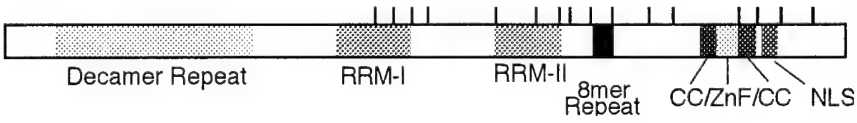
may mediate the suppression of A9 tumorigenesis. Whereas parental A9 cells undergo apoptosis in response to chemotherapeutic drugs, SEMA3F expressing clones do not and are instead growth arrested. It is also very interesting that the receptors for SEMA3F include neuropilins 1 and 2, apparently in concert with plexins. Importantly, recent studies have shown that the angiogenic factor, $VEGF_{165}$, also binds the neuropilins. Thus there is the liklihood of complex inhibitory interactions between this secreted semaphorin and $VEGF_{165}$, in the regulation of angiogenesis and perhaps cell migration. SEMA3F may indeed be a critical gene in breast tumors, despite the absence of clear deletions in these tumors.

Our studies on the semaphorin molecule have included the development of polyclonal antibodies specific for SEMA3F. Because of the high degree of conservation among the family members, this was not a trivial task. While these antibodies (directed against a unique peptide) appear to work well for immunohistochemistry, they do not work on Western blots. We have found that the normal membrane staining for SEMA3F is lost and that this correlates inversely with staining for VEGF. This appears to be a very fruitful area for investigation which we plan on continuing. We generated mammalian expression constructs for SEMA3F and introduced these into mouse A9 cells. Unlike the results from the Naylor lab (Naylor et al., 1998), we observed no difference in tumorigenicity. However, while expression of the recombinant gene was clearly evident at the RNA level, we have been unable to clearly demonstrate the presence of recombinant SEMA3F protein. The difficulties included; 1) the mouse and human genes are very identical, 2) SEMA3F is a secreted molecule with a signal sequence at its amino terminus such that epitope tags cannot be placed in this region, 3) the carboxy terminus of the semaphorins have been reported to undergo furin mediated cleavage making this site problematic for an epitope tag. Analysis of the transfected A9 cells using our polyclonal antibody indicated that a small subgroup express the human protein at high levels. We think this is due to the variable copy number of the episomal vector used for the construct.

6.C.4 Investigation of DEF-3(NY-LU-12/g16)

We noted that each of the initial three homozygous deletions involving 3p21.31 appeared to be clustered at their telomeric boundaries. Therefore, we looked for a gene which might be commonly interrupted by these alterations. DEF-3 is a novel sequence which was isolated originally by us as a partially processed cDNA (L1-204) containing a single, very small exon (Roche et al., 1996). We subsequently cloned the full length cDNA from both human and mouse and found that the predicted protein has significant homologies to RNA binding proteins (see Figure below) with multiple domains for RNA binding and protein-protein interactions.

hDEF-3 (g16/NY-LU-12).



Domains of hDEF-3. RNA recognition motifs (RRM), octamer repeat (8mer), coiled coil (CC), zinc finger (ZnF) and nuclear localization signal (NLS) are indicated.

RNA binding proteins are involved in a diversity of biological functions affecting RNA processing, half-lives and expressibility. Interestingly, antibodies against RNA binding proteins have been identified in some breast cancer patients with the POMA syndrome resulting from antibodies directed against the RNA binding protein Nova-1 (Buckanovich et al., 1996). Gure et al., (Gure et al., 1998) have independently identified antibodies against DEF-3/NY-LU-12 in some patients with lung cancer. These observations raise the possibility that antibodies against DEF-3 could be related to certain deletions in 3p21. Our efforts on this gene have led us to define a new family of RNA binding proteins and to demonstrate that DEF-3 and an adjacent, co-deleted, family member (LUCA15) bind poly (G) homopolymers *in vitro* (see Figure 5 in Drabkin et al., (1999) in Appendix). Moreover, the RNA recognition domains of DEF-3 and LUCA15 are non-reactive with antibodies from patients with the Hu syndrome. Anti-Hu sera were tested because the RRMs of DEF-3 and LUCA15 are similar to those from the Hu proteins. We have developed antibodies

against the RRM domains of DEF-3 which are undergoing characterization and which would permit immunologic detection of deletions in direct tumor materials and facilitate analysis of this region.

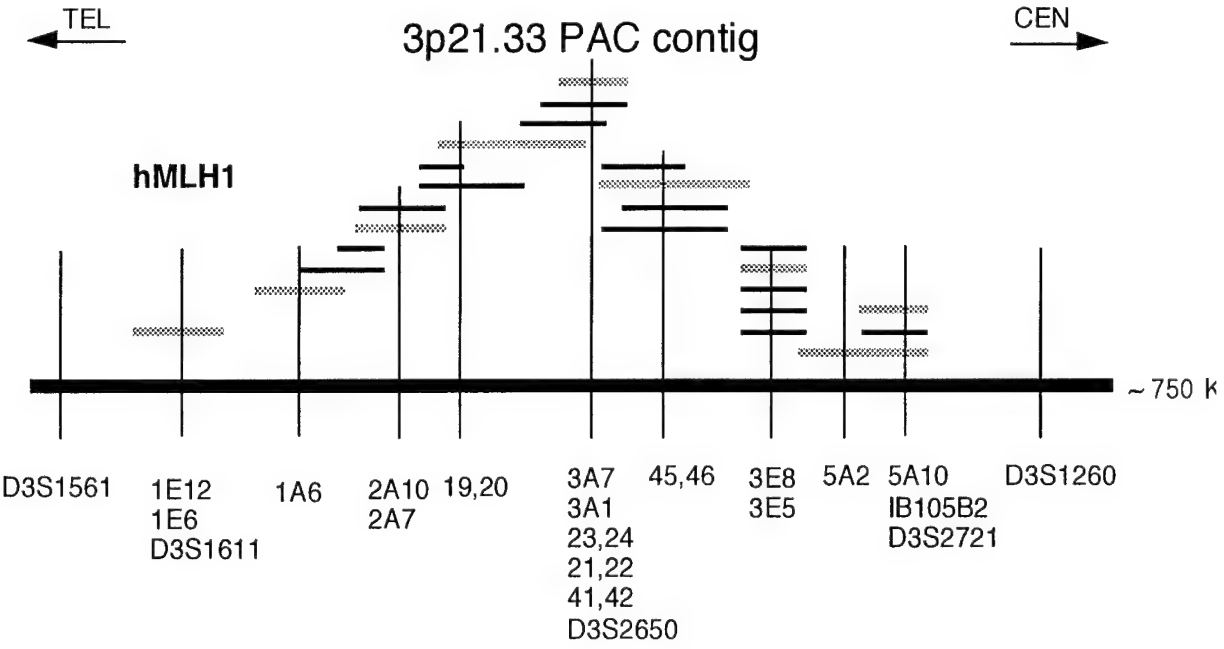
6.C.5. Summary and Significance of 3p21.3 Findings

The 3p21.3 deletion region has been an important region for the identification of a tumor suppressor gene. One of the best candidates to date is the SEMA3F gene, cloned in our laboratory. The DEF-3 gene is also a candidate. Future experimental directions will focus on the function of SEMA3F since it can directly interferre with binding of neuropilin receptors by $VEGF_{165}$. SEMA3F therefore has great potential significance for tumor angiogenesis and possibly other functions of VEGF including cell migration and growth. Data obtained for this molecule will likely have significance not only for breast cancers but for many other carcinomas as well.

6.D. Involvement of 3p21.33 Region in Breast Carcinoma

6.D.1. Screening Breast Tumors for Deletions in the 3p21.33 Deletion Region.

Chromosomal band 3p21 harbors a second homozygous deletion region, originally described by Yamakawa et al., (Yamakawa et al., 1993). They went on to show that the minimal deletion spanned approximately 800 kb and was covered by YAC 936C1. We demonstrated that the deletion target was separated from the SEMA3F deletion site by 10 to 15 megabases of DNA (Roche et al., 1996). In the course of mapping this region more thoroughly, we discovered that one of our markers, Mbo16E2, was present in this YAC (Roche et al., 1996). In previous Progress Reports, we described how the YAC 936C1 was found to be chimeric and thus useless for developing additional markers in this interval. We proceeded to generate an independent clone contig using P1 artificial chromosomes (PACs) which covers at least 750 kb (see Figure below).



PAC clone contig covering the 3p21.33 deletion region. Clones are symbolized by short horizontal lines, the shaded lines represent PACs which were sub-cloned and used to generate markers. A sub-set of the markers utilized for screening breast tumors are listed below the chromosome (thick solid line) with their approximate location. The contig is estimated to span 600 to 750 kb.

We have prepared sub-clone libraries from many of these PACs, sequenced selected clones and developed PCR primer pairs suitable for deletion analysis. Twenty six STS markers from this region have been used to screen 12 breast tumor cell lines. These markers included specific primer pairs for two genes;

the mismatch repair gene hMLH1 and Integrin-α_{RLC} . The screen showed that none of these breast tumors harbored deletions of this region. The development of this contig and the results of our screening assays are currently in preparation for publication.

6.D.2 Significance

This region was one of the initial target deletions proposed in our application. The reagents we developed will be useful for ongoing investigations into potential alterations in this.

6.E. Exploration of developmental signaling pathways in cancer.

Our discovery of the TRC8 gene and elucidation of its structure led us to investigate the relationship between developmental signaling pathways and cancer. In particular, we noted the relationship between the Hedgehog and WNT pathways, and the relationship between HOX gene expression, WNT regulation of HOX and the downstream targets for HOX proteins, particularly the FGF genes. These considerations led us to survey a series of carcinoma cell lines for expression patterns of genes which make up these pathways. While these studies have been beyond the scope of the present project, they were stimulated by our results which directly resulted from the Army support, namely the discovery of the TRC8 gene. Importantly, some of these signaling systems, particularly the WNT pathway, have been previously implicated in breast cancer. These studies have led to two manuscripts which are currently undergoing review. Importantly, we are in the process of developing several technological approaches for detecting expression of key signaling molecules in these pathways. These methods are based upon cDNA array hybridization and quantitative real time PCR analyses. While our Army supported project is now completed, we are continuing to pursue these observations through other mechanisms. In particular, an application is now pending at NIH for a pilot study looking at the possible utility of using expression data as a means for identifying and discriminating among breast tumors.

7. KEY RESEARCH ACCOMPLISHMENTS.

- Discovered the homozygous deletion region in 3p14 which affects breast tumors.
- Cloned and sequenced the deletion region, revealing it is gene poor.
- Demonstrated that deletions coincided with FRA3B.
- Demonstrated that the region is genetically unstable in selected tumors.
- Discovered the TRC8 gene, suggesting a new pathway to cancer.
- Established parameters of expression and function for the SEMA3F gene, likely to influence VEGF signaling.
- Discovered the DEF-3(g16/NY-LU-12) gene and demonstrated that it encodes a developmentally regulated RNA binding protein.
- Developed a PAC clone contig for the 3p21.33 homozygous deletion region but found it was not affected in lung tumors tested.

8. REPORTABLE OUTCOMES.

Our studies have led directly to 6 publications with at least 2 more in preparation. In addition, these efforts have been reported as abstracts for a number of meetings (Cold Spring Harbor, DOD Era of Hope and American Society of Human Genetics) and have been incorporated into numerous presentations. Importantly, one funded grant application has been a direct result of this project while several others have been submitted. Two of these are pending with NIH and a third will be submitted soon. The publications are listed below:

Publications:

Boldog, F., Gemmill, R.M., West, J., Robinson, M., Robinson, L., Li, E., Roche, J., Todd, S., Lundstrom, R., Jacobson, J., Klinger, H., and Drabkin, H.A. (1997). Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B. Hum. Mol. Genet. *6,* 193-203.

Le Beau MM, Drabkin H, Glover TW, Gemmill R, Rassool FV, McKeithan TW and Smith DI. (1998). A *FHIT* tumor suppressor gene. Genes, Chromosomes and Cancer 21:281-289.

Roche J, West J, Gemmill R et Harry Drabkin. (1998).Chromosome 3p et recherche de gènes suppresseurs de tumeurs et gènes de sémaphorines en 3p21.3. Médecine et Sciences14:283-290.

Gemmill RM, West JD, Boldog F, Tanaka N, Robinson LJ, Smith DI, Li F and Drabkin HA (1998). The hereditary renal cell carcinoma (RCC) 3;8 translocation fuses FHIT to a novel patched related gene, TRC8. PNAS 95:9572-9577.

Drabkin HA, West JD, Hotfilder M, Heng YM, Erickson P, Calvo R, Dalmau J, Gemmill RM and SablitzkyF (1999). DEF-3(g16/NY-LU-12), An RNA Binding Protein from the 3p21.3 Homozygous Deletion Region in SCLC. Oncogene 18: 2589-2597.

F de Castro, L Hu, HA Drabkin, C Sotelo and Alain Chédotal. Chemoattraction and Chemorepulsion of olfactory bulb axons by different secreted semaphorins (1999). J. Neuroscience 19(11): 4428-4436.

Grant Applications:

NIH, 1RO1 CA76035-01A1, "Molecular-Genetic Analysis of 3p14 Genomic Stability"
NIH, R01 (PENDING, 1999) "Genetic Characterization of dTRC8"
NIH, R21 (PENDING, 1999) "WNT-HOX-FGF Pathway Gene Expression in Breast CA"

9. CONCLUSIONS

In each of the previous sections comprising the Body of this report, we have provided a summary of our important results and their significance; the most important of these will be recapitulated here. We identified one chromosome 3p region (3p14.2) subject to homozygous deletion in breast cancers, but our best evidence to date suggests that these deletions are the result of inherent genomic instability. We generated strong evidence that the postulated tumor suppressor gene, FHIT, is not the target of these alterations. We are continuing to explore the observation of genomic instability, attempting to define the signaling pathways which influence it. Importantly, our exploration of 3p14 led directly to the isolation of the TRC8 gene which may define a new pathway to tumorigenesis. Of equal significance, this discovery led us to explore the role in cancer of other developmental signaling modalities, including the WNT, HOX and FGF pathways. Our initial results suggest that differential patterns of expression for selected genes in these pathways may be able to identify tumor types and/or discriminate tumors with distinct biological properties. Two other frequently deleted sites were surveyed by us, one of which had been previously reported with homozygous deletions in breast cancers. Our studies uncovered a number of genes whose locations in these deletion targets suggest possible roles in tumor development. Two candidate genes found in the proximal 3p21.3 deletion region, SEMA3F and DEF-3, are of particular interest; expression constructs and antibodies have been generated for both. Importantly, what little is known about the functions of these genes suggests they will have biological roles which are broadly significant to the development of many cancers, including those of breast. SEMA3F, which we originally cloned, may be the long sought after 3p21 tumor suppressor gene. We predicted and confirmed that DEF-3 was an RNA binding protein and it could have roles in cellular differentiation, similar to other RNA binding proteins. TRC8 has a structure suggestive of a signaling receptor and our initial experiments support a role in early development. All these biological functions are in great need of exploration and we are committed to continuing these investigations.

What do these results mean to the "person on the street"? Actually, quite a lot. The development of breast cancer is a complex process involving alterations in many genes and cellular pathways. While we know of some pathways altered in breast cancer (eg., Her2/neu, estrogen receptor, etc.) there are many more which remain unknown. Our efforts are focused on uncovering more of these unknown routes to

breast cancer. With such knowledge, it should be possible to devise much more effective treatments by targeting specific pathways or combinations of pathways.

10. BIBLIOGRAPHY.

PUBLICATIONS RESULTING FROM THIS PROJECT:

Boldog, F., Gemmill, R.M., West, J., Robinson, M., Robinson, L., Li, E., Roche, J., Todd, S., Lundstrom, R., Jacobson, J., Klinger, H., and Drabkin, H.A. (1997). Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B. Hum. Mol. Genet. *6,* 193-203.

Le Beau MM, Drabkin H, Glover TW, Gemmill R, Rassool FV, McKeithan TW and Smith DI. (1998). A *FHIT* tumor suppressor gene. Genes, Chromosomes and Cancer 21:281-289.

Roche J, West J, Gemmill R et Harry Drabkin. (1998).Chromosome 3p et recherche de gènes suppresseurs de tumeurs et gènes de sémaphorines en 3p21.3. Médecine et Sciences14:283-290.

Gemmill RM, West JD, Boldog F, Tanaka N, Robinson LJ, Smith DI, Li F and Drabkin HA (1998). The hereditary renal cell carcinoma (RCC) 3;8 translocation fuses FHIT to a novel patched related gene, TRC8. PNAS 95:9572-9577.

Drabkin HA, West JD, Hotfilder M, Heng YM, Erickson P, Calvo R, Dalmau J, Gemmill RM and SablitzkyF (1999). DEF-3(g16/NY-LU-12), An RNA Binding Protein from the 3p21.3 Homozygous Deletion Region in SCLC. Oncogene 18: 2589-2597.

F de Castro, L Hu, HA Drabkin, C Sotelo and Alain Chédotal. Chemoattraction and Chemorepulsion of olfactory bulb axons by different secreted semaphorins (1999). J. Neuroscience 19(11): 4428-4436.

ABSTRACTS RESULTING FROM THIS PROJECT:

Naylor SL, Hensel CH, Xiang RH, Daly MC, Garcia D, Buys CHCM, Drabkin H, Carlson H, Kok K, and Kerbacher K. Tumor suppresion and the 21.3 region of chromosome 3. Am. Journal of Hum Genet. 57 (supp): A73, 1995

Drabkin H, Boldog F, Robinson M, Roche J, Giardini K, Lundstrom R, Robinson L and Gemmill RM. A Recurrent Homozygously Deleted Region in Human Carcinoma Cells Affecting 3p14. Proc. AACR 37: pg. 127 #3515, 1996 .

H. Drabkin F. Boldog, J. Roche, M. Robinson, L. Robinson W. Franklin, H. Klinger and R. Gemmill. Homozygous deletions affecting 3p14.2 in human carcinoma cell lines: A case of genomic instability? Cancer Genetics & Tumor Suppressor Genes. CSH Meeting, Aug 8-14, 1996.

H. Drabkin, M Robinson, R Lundstrom, J Cypser, E Li and R Gemmill. Chromosome 3 beyond the physical map. Brazilian Journal of Genetics 19, No 2 Supp. #158, pg. 131, 1996 (9th International Congress of Human Genetics).

F. Boldog, R. Gemmill, J. West, M. Robinson, L. Robinson, E. Li, J. Roche, S. Todd, J. Jacobsen, H. Klinger and H. Drabkin. 3p14 Homozygous Deletions in Carcinomas of Lung, Breast, Colon and Cervix and Sequence Analysis of FRA3B. 8th World Conference on Lung Cancer. Lung Cancer 18: 141, 1997.

Todd S, Franklin WA, Varella-Garcia M, Kennedy T, Hilliker CE, Hahner L, Anderson M, Weist JS, Drabkin HA and Gemmill RM. Homozygous deletions of human chromosome 3p in lung tumors. Lung Cancer 18: 14157, 1997.

Boldog F, J West, R Gemmill and H Drabkin (1997). 3p Deletions in Breast Cancer. Presented at the Era of Hope Breast Cancer Research Meeting, DOD, Washington DC, Oct 31-Nov 4, 1997.

RM Gemmill, JD West, F Boldog, N Tanaka, LJ Robinson, DI Smith, F Li and HA Drabkin (1998). The hereditary renal cell carcinoma 3;8 translocation fuses FHIT to a novel *patched* related gene, TRC8. presented: Cold Spring Harbor, Pathways to Cancer March 11-14, 1998.

RM Gemmill, JD West, F Boldog, N Tanaka, LJ Robinson, DI Smith, FP Li and HA Drabkin (1998) The hereditary renal cell cancer 3;8 translocation fuses FHIT to a novel *patched* related gene, TRC8. Am J. Hum. Genet. 63:pgA38, #201. Presented at the 48th annual meeting of the American Society of Human Genetics, Oct 27-31, 1998.

PERSONNEL SUPPORTED BY THIS PROJECT:

Year 1 1994-1995
Kathy Bodor
Ferenc Boldog, Ph.D.
Harry A. Drabkin, M.D.
Misi Robinson

Year 2 1995-1996
Ferenc Boldog, Ph.D.
Harry Drabkin, M.D.
Robert Gemmill, Ph.D.
Jan Jacobsen
Linda Robinson

Year 3 1996-1997
Ferenc Boldog, Ph.D.
Harry Drabkin, M.D.
Robert Gemmill, Ph.D.
Jan Jacobsen
Linda Robinson

Year 4 1997-1998
Harry Drabkin, M.D.
Paul Erickson, Ph.D.
Robert Gemmill, Ph.D.
Linda Robinson
James West, Ph.D.

Year 5 1998-1999 Time Extension, No Money
Harry Drabkin, M.D.
Paul Erickson, Ph.D.
Robert Gemmill, Ph.D.

11. REFERENCES.

Ahmadian, M., Wistuba, I.I., Fong, K.M., Behrens, C., Kodagoda, D.R., Saboorian, M.H., Shay, J., Tomlinson, G.E., Blum, J., Minna, J.D., and Gazdar, A.F. (1997). Analysis of the FHIT gene and FRA3B region in sporadic breast cancer, preneoplastic lesions, and familial breast cancer probands. Cancer Res. *57*, 3664-3668.

Ali, I.U., Lidereau, R., and Callahan, R. (1989). Presence of two members of c-erbA receptor gene family (c-erbA beta and c-erbA2) in smallest region of somatic homozygosity on chromosome 3p21-p25 in human breast carcinoma. J. Natl. Cancer. Inst. *81*, 1815-1820.

Bergthorsson, J.T., Eiriksdottir, G., Barkardottir, R.B., Egilsson, V., Arason, A., and Ingvarsson, S. (1995). Linkage analysis and allelic imbalance in human breast cancer kindreds using microsatellite markers from the short arm of chromosome 3. Hum. Genet *96*, 437-443.

Boldog, F., Gemmill, R.M., West, J., Robinson, M., Robinson, L., Li, E., Roche, J., Todd, S., Waggoner, B., Lundstrom, R., Jacobson, J., Mullokandov, M.R., Klinger, H., and Drabkin, H.A. (1997). Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B. Hum. Mol. Genet. *6*, 193-203.

Borodovsky, M. and McIninch, J. (1993). GenMark: Parallel gene recognition for both DNA strands. Computers Chem. *17*, 123-133.

Buchhagen, D.L., Qiu, L., and Etkind, P. (1994). Homozygous deletion, rearrangement and hypermethylation implicate chromosome region 3p14.3-3p21.3 in sporadic breast-cancer development. Int. J Cancer *57*, 473-479.

Buckanovich, R.J., Yang, Y.Y., and Darnell, R.B. (1996). The onconeural antigen Nova-1 is a neuron-specific RNA-binding protein, the activity of which is inhibited by paraneoplastic antibodies. J Neurosci. *16*, 1114-1122.

Chen, L.C., Matsumura, K., Deng, G., Kurisu, W., Ljung, B.M., Lerman, M.I., Waldman, F.M., and Smith, H.S. (1994). Deletion of two separate regions on chromosome 3p in breast cancers. Cancer Res. *54*, 3021-3024.

Cohen, A.J., Li, F.P., Berg, S., Marchetto, D.J., Tsai, S., Jacobs, S.C., and Brown, R.S. (1979). Hereditary renal-cell carcinoma associated with a chromosomal translocation. N. Engl. J. Med. *301*, 592-595.

Devilee, P., van den Broėk, M., Kuipers-Dijkshoorn, N., Kolluri, R., Khan, P.M., Pearson, P.L., and Cornelisse, C.J. (1989). At least four different chromosomal regions are involved in loss of heterozygosity in human breast carcinoma. Genomics. *5*, 554-560.

Drabkin, H.A., West, J.D., Hotfilder, M., Heng, Y.M., Erickson, P., Calvo, R., Dalmau, J., Gemmill, R.M., and Sablitzky, F. (1999). DEF-3(g16/NY-LU-12), an RNA binding protein from the 3p21.3 homozygous deletion region in SCLC. Oncogene *In Press*,

Fazeli, A., Dickinson, S.L., Hermiston, M.L., Tighe, R.V., Steen, R.G., Small, C.G., Stoeckli, E.T., Keino-Masu, K., Masu, M., Rayburn, H., Simons, J., Bronson, R.T., Gordon, J.I., Tessier-Lavigne, M., and Weinberg, R.A. (1997). Phenotype of mice lacking functional Deleted in colorectal cancer (Dcc) gene. Nature *386*, 796-804.

Gemmill, R.M., Varella-Garcia, M., Smith, D.I., Erickson, P., Golembieski, W., Miller, Y., Coyle-Morris, J., Tommerup, N., and Drabkin, H.A. (1991). A 2.5-Mb physical map within 3p21.1 spans the breakpoint associated with Greig cephalopolysyndactyly syndrome. Genomics. *11*, 93-102.

Gemmill, R.M., West, J.D., Boldog, F.L., Tanaka, N., Robinson, L.J., Smith, D.I., Li, F., and Drabkin, H.A. (1998). The hereditary renal cancer 3;8 translocation fuses FHIT to a *patched*-related gene, *TRC8*. Proc. Natl. Acad. Sci. U. S. A. *95*, 9572-9577.

Geurts, J.M., Schoenmakers, E.F., Roijer, E., Stenman, G., and Van de Ven, W.J. (1997). Expression of reciprocal hybrid transcripts of HMGIC and FHIT in a pleomorphic adenoma of the parotid gland. Cancer Res. *57*, 13-17.

Glover, T.W., Berger, C., Coyle, J., and Echo, B. (1984). DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. Hum. Genet. *67*, 136-142.

Gure, AO., Altorki, NK., Stockert, E., Scanlan, MJ., Old, LJ., and Chen, Y-T. (1998). Human Lung Cancer Antigens Recognized by Autologous Antibodies: Definition of a Novel cDNA Derived from the Tumor Suppressor Gene Locus on Chromosome 3p21.3. Cancer Research *58*, 1034-1041.

Hainsworth, P.J., Raphael, K.L., Stillwell, R.G., Bennett, R.C., and Garson, O.M. (1991). Cytogenetic features of twenty-six primary breast cancers. Cancer. Genet. Cytogenet. *53*, 205-218.

Inoue, H., Ishii, H., Alder, H., Snyder, E., Druck, T., Huebner, K., and Croce, CM. (1997). Sequence of the FRA3B common fragile region: implications for the mechanism of FHIT deletion. Proc. Natl. Acad. Sci. U. S. A. *94(26)*, 14584-14589.

Ji, L., Fang, B., Yen, N., Fong, K., Minna, J., and Roth, J.A. (1999). Induction of apoptosis and inhibition of tumorigenicity and tumor growth by adenovirus vector-mediated fragile histidine triad (FHIT) gene overexpression. Cancer Res. *59*, 3333-3339.

Keino-Masu, K., Masu, M., Hinck, L., Leonardo, E.D., Chan, S.S., Culotti, J.G., and Tessier-Lavigne, M. (1996). Deleted in Colorectal Cancer (DCC) encodes a netrin receptor. Cell *87*, 175-185.

Killary, A.M., Wolf, M.E., Giambernardi, T.A., and Naylor, S.L. (1992). Definition of a tumor suppressor locus within human chromosome 3p21-p22. Proc. Natl. Acad. Sci. U. S. A. *89*, 10877-10881.

Livingstone, L.R., White, A., Sprouse, J., Livanos, E., Jacks, T., and Tlsty, T.D. (1992). Altered cell cycle arrest and gene amplification potential accompany loss of wild-type p53. Cell *70*, 923-935.

Lowe, S.W., Ruley, H.E., Jacks, T., and Housman, D.E. (1993). p53-dependent apoptosis modulates the cytotoxicity of anticancer agents. Cell *74*, 957-967.

Lu, Y.J., Xiao, S., Yan, Y.S., Fu, S.B., Liu, Q.Z., and Li, P. (1993). Direct chromosome analysis of 50 primary breast carcinomas. Cancer Genet Cytogenet. *69*, 91-99.

Lucito, R., Nakimura, M., West, J.A., Han, Y., Chin, K., Jensen, K., McCombie, R., Gray, J.W., and Wigler, M. (1998). Genetic analysis using genomic representations. Proc. Natl. Acad. Sci. U. S. A. *95*, 4487-4492.

Nancarrow, J.K., Holman, K., Mangelsdorf, M., Hori, T., Denton, M., Sutherland, G.R., and Richards, R.I. (1995). Molecular basis of p(CCG)n repeat instability at the FRA16A fragile site locus. Hum. Mol. Genet *4*, 367-372.

Naylor, S.L., Davalos, A.R., Hensel, C.H., and Xiang, R.H. (1998). Human semaphorin IV gene from 3p21.3 suppresses tumor growth in nude mice and attenuates apoptosis. Am. J. Human. Genet. *63*, A80.

Negrini, M., Monaco, C., Vorechovsky, I., Ohta, M., Druck, T., Baffa, R., Huebner, K., and Croce, C.M. (1996). The FHIT gene at 3p14.2 is abnormal in breast carcinomas. Cancer Res. *56*, 3173-3179.

Ohta, M., Inoue, H., Cotticelli, M.G., Kastury, K., Baffa, R., Palazzo, J., Siprashvili, Z., Mori, M., McCue, P., Druck, T., Croce, C.M., and Huebner, K. (1996). The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. Cell *84*, 587-597.

Panagopoulos, I., Pandis, N., Thelin, S., Petersson, C., Mertens, F., Borg, A., Kristoffersson, U., Mitelman, F., and Aman, P. (1996). The FHIT and PTPRG genes are deleted in benign proliferative breast disease associated with familial breast cancer and cytogenetic rearrangements of chromosome band 3p14. Cancer Res. *56*, 4871-4875.

Pandis, N., Jin, Y., Limon, J., Bardi, G., Idvall, I., Mandahl, N., Mitelman, F., and Heim, S. (1993). Interstitial deletion of the short arm of chromosome 3 as a primary chromosome abnormality in carcinomas of the breast. Genes. Chromosom. Cancer. *6*, 151-155.

Pintado, E., de Diego, Y., Hmadcha, A., Carrasco, M., Sierra, J., and Lucas, M. (1995). Instability of the CGG repeat at the FRAXA locus and variable phenotypic expression in a large fragile X pedigree. J Med. Genet *32*, 907-908.

Roche, J., Boldog, F., Robinson, M., Robinson, L., Varella-Garcia, M., Swanton, M., Waggoner, B., Fishel, R., Franklin, W., Gemmill, R., and Drabkin, H. (1996). Distinct 3p21.3 deletions in lung cancer and identification of a new human semaphorin. Oncogene. *12*, 1289-1297.

Sato, T., Akiyama, F., Sakamoto, G., Kasumi, F., and Nakamura, Y. (1991). Accumulation of genetic alterations and progression of primary breast cancer. Cancer. Res. *51*, 5794-5799.

Scully, R., Chen, J., Ochs, R.L., Keegan, K., Hoekstra, M., Feunteun, J., and Livingston, D.M. (1997a). Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. Cell *90*, 425-435.

Scully, R., Chen, J., Plug, A., Xiao, Y., Weaver, D., Feunteun, J., Ashley, T., and Livingston, D.M. (1997b). Association of BRCA1 with Rad51 in mitotic and meiotic cells. Cell *88*, 265-275.

Sekido, Y., Bader, S., Latif, F., Chen, J.Y., Duh, F.M., Wei, M.H., Albanesi, J.P., Lee, C.C., Lerman, M.I., and Minna, J.D. (1996). Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung cancer deletion region and demonstrate distinct expression patterns. Proc. Natl. Acad. Sci. U. S. A. *93*, 4120-4125.

Sekido, Y., Ahmadian, M., Wistuba, I.I., Latif, F., Bader, S., Wei, M.H., Duh, F.M., Gazdar, A.F., Lerman, M.I., and Minna, J.D. (1998). Cloning of a breast cancer homozygous deletion junction narrows the region of search for a 3p21.3 tumor suppressor gene. Oncogene. *16*, 3151-3157.

Teixeira, M.R., Pandis, N., Bardi, G., Andersen, J.A., and Heim, S. (1996). Karyotypic comparisons of multiple tumorous and macroscopically normal surrounding tissue samples from patients with breast cancer. Cancer Res. *56*, 855-859.

Todd, M.C., Xiang, R.H., Garcia, D.K., Kerbacher, K.E., Moore, S.L., Hensel, C.H., Liu, P., Siciliano, M.J., Kok, K., van den Berg, A., Veldhuis, P., Buys, C.H., Killary, A.M., and Naylor, S.L. (1996). An 80 Kb P1 clone from chromosome 3p21.3 suppresses tumor growth in vivo. Oncogene. *13*, 2387-2396.

Xiang, R.H., Hensel, C.H., Garcia, D.K., Carlson, H.C., Kok, K., Daly, M.C., Kerbacher, K., van den Berg, A., Veldhuis, P., Buys, C.H., and Naylor, S.L. (1996). Isolation of the human semaphorin III/F gene (SEMA3F) at chromosome 3p21, a region deleted in lung cancer. Genomics. *32*, 39-48.

Yamada, T., Endo, R., Gotoh, M., and Hirohashi, S. (1997). Identification of semaphorin E as a non-MDR drug resistance gene of human cancers. Proc. Natl. Acad. Sci. U. S. A. *94*, 14713-14718.

Yamakawa, K., Takahashi, T., Horio, Y., Murata, Y., Takahashi, E., Hibi, K., Yokoyama, S., Ueda, R., and Nakamura, Y. (1993). Frequent homozygous deletions in lung cancer cell lines detected by a DNA marker located at 3p21.3-p22. Oncogene. *8*, 327-330.

12. APPENDIX.

Copies of publications resulting from this project are included here.

Human Molecular Genetics, 1997, Vol. 6, No. 2 193–203

ARTICLE

Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B

Ferenc Boldog[1,‡], Robert M. Gemmill[1,‡], James West[1], Misi Robinson[1], Linda Robinson[1], Efang Li[1], Joelle Roche[1,+], Sean Todd[1], Barbara Waggoner[1], Ron Lundstrom[2], Jan Jacobson[1], Michael R. Mullokandov[3], Harold Klinger[3] and Harry A. Drabkin[1,*]

[1]Division of Medical Oncology, University of Colorado Health Sciences Center, 4200 E. 9th Ave., Denver, CO 80262, USA, [2]Eleanor Roosevelt Institute, 1899 Gaylord, Denver, CO 80206, USA and [3]Department of Molecular Genetics, Albert Einstein College of Medicine, 1300 Morris Park Ave., New York, NY 10461, USA

Received September 16, 1996; Revised and Accepted November 27, 1996

Loss of heterozygosity (LOH) involving 3p occurs in many carcinomas but is complicated by the identification of four distinct homozygous deletion regions. One putative target, 3p14.2, contains the common fragile site, FRA3B, a hereditary renal carcinoma-associated 3;8 translocation and the candidate tumor suppressor gene, *FHIT*. Using a ~300 kb cosmid/λ contig, we identified homozygous deletions in cervix, breast, lung and colorectal carcinoma cell lines. The smallest deletion (CC19) was shown not to involve *FHIT* coding exons and no DNA sequence alterations were present in the transcript. We also detected discontinuous deletions as well as deletions in non-tumor DNAs, suggesting that *FHIT* is not a selective target. Further, we demonstrate that some reported *FHIT* aberrations represent normal splicing variation. DNA sequence analysis of 110 kb demonstrated that the region is high in A–T content, LINEs and MER repeats, whereas Alu elements are reduced. We note an intriguing similarity in repeat sequence composition between FRA3B and a 152 kb segment from the Fragile-X region. We also identified similarity between a FRA3B segment and a small polydispersed circular DNA. In contrast to the selective loss of a tumor suppressor gene, we propose an alternative hypothesis, that some putative targets including FRA3B may undergo loss as a consequence of genomic instability. This instability is not due to DNA mismatch repair deficiency, but may correlate in part with p53 inactivation.

INTRODUCTION

LOH involving 3p occurs frequently in carcinomas of the lung, kidney, cervix, breast and other epithelial neoplasms (1–7). However, 3p loss is complex, involving at least four distinct homozygously deleted regions (8–12). One of the most frequently lost regions is 3p14, especially in cervical carcinomas (13). This region is also of interest since it contains the site of a hereditary renal carcinoma-associated translocation, t(3;8)(p14.2;q24.1) (14), and is the location of the most inducible common fragile site in the genome, FRA3B (15). We previously reported cloning of the 3;8 translocation breakpoint (16) and demonstrated by fluorescence *in situ* hybridization that FRA3B was further telomeric (17). These studies also suggested that breaks in FRA3B occur over a region instead of at a single site. While searching for genes in this region, we identified a homozygous deletion in the cervical carcinoma cell line, HeLa, involving marker *D3S1300*.

We developed a ~300 kb cosmid/λ contig within FRA3B containing *D3S1300*. Probes from the region detected frequent homozygous deletions in cervical, lung, colorectal and breast carcinoma cell lines. Cervical carcinomas, which are associated with papilloma virus infection and p53 inactivation (18), were most frequently deleted and the smallest deletion occurred in cell line CC19. During our investigations, Ohta *et al.* reported identification of a candidate tumor suppressor gene, *FHIT*, which spanned the t(3;8) breakpoint and was deleted in various carcinoma cell lines (19). However, our analysis indicates that *FHIT* is not the target of these deletions. We also observed that 3p14 deletions tend not to occur in tumors with deficiencies in DNA mismatch repair.

*To whom correspondence should be addressed
+Present address: CNRS URA1172, IBMIG, Universite de Poitiers, 40 Av. du Recteur Pineau, 86022 Poitiers Cédex, France
‡These authors contributed equally to this work

Table 1. Results of microsatellite analysis and 3p14 deletions in selected cell lines

Cell line	Mismatch repair deficiency/ microsatellite instability[a]	3p14 status	Reference
Colorectal carcinoma[b]			
HT-29	intact	deletion	this study
SW-480	intact	deletion	Umar *et al.* (25); this study
SW-403	intact	deletion	this study
SW-948	intact	rearranged	thisstudy
SW-48	defective	no deletion	Boyer *et al.* (23); this study
DLD1/HCT15	defective	no deletion	da Costa *et al.* (24)
		Bhattacharyya *et al.* (22)	
Cervical carcinoma[c]			
C-41	intact	deletion	Larson *et al.* (26)[d]
Caski	intact	deletion	Larson *et al.* (26)[d]
SIHA	intact	deletion	Larson *et al.* (26)[d]
HeLa	intact	deletion	Boyer *et al.* (23); Umar *et al.* (25)
CC19	NT	deletion	
MS-751	intact	deletion	Larson *et al.* (26)[d]
C-33A	defective	deletion	Larson *et al.* (26)[d]; this study
ME-180	intact	no deletion	Larson *et al.* (26)[d]

[a]Defective refers to a reported deficiency in mismatch repair or microsatellite instability.
[b]Overall, deletions (5) or rearrangements (1) were identified in 6/12 colorectal carcinoma DNAs. In addition to those listed, deletions occurred in COLO-205 and COLO-320. No deletions were identified in SKCO-1, CaCo-2, SW-1417 and T-84.
[c]Represents all cervical carcinoma samples examined.
[d]Personal communication from Dr Garret M. Hampton.
NT = not tested.

DNA sequencing studies were performed to identify features that might provide insight into these breaks. Our results demonstrate that FRA3B is high in A–T content, LINES and MER repeats. In contrast, Alu elements and confirmed genes are reduced. We identified a FRA3B segment highly similar to a reported small polydispersed circular DNA, sequences which are markedly elevated in damaged or unstable genomes. FRA3B also bears an intriguing overall sequence similarity to the Fragile-X region. However, unlike the rare folate-sensitive fragile sites, no triplet repeats nor methylated CpG island was identified. These overall features may be responsible for, or contribute to, the observed instability of this region.

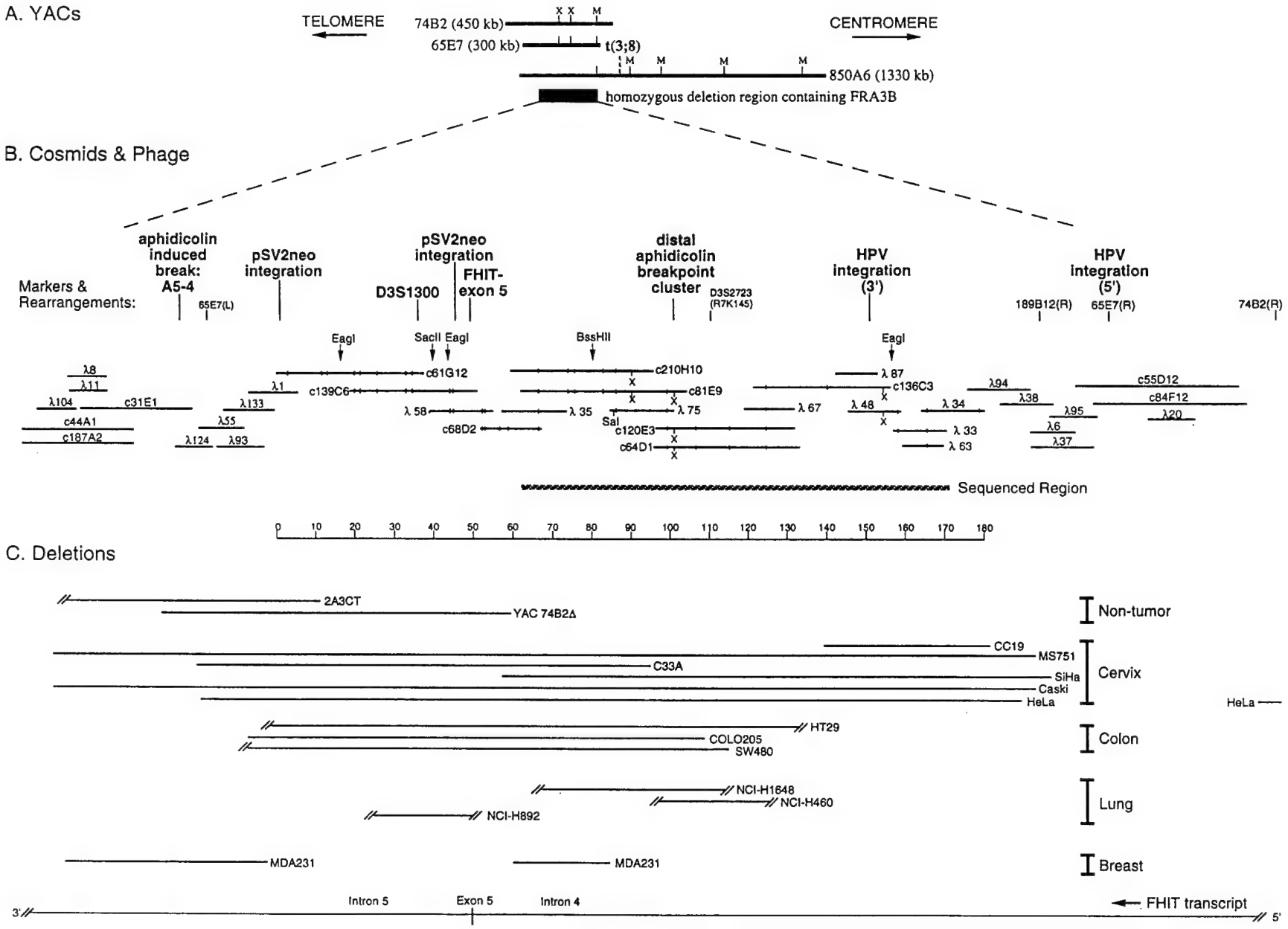
RESULTS

Identification of homozygous deletion

Following identification of the HeLa deletion, additional cervical carcinoma lines were examined using probes spanning a ~300 kb cosmid/phage contig from the FRA3B region (Fig. 1A and B). Figure 2A shows an example using cosmid c136C3. Deletions were evident in DNAs from MS751, SIHA and Caski (lanes 2, 5 and 7) whereas CC19 (lane 6) shows missing and altered bands (arrow). Altogether, homozygous deletions were detected in seven out of eight lines (87.5%). Similar hybridizations were performed using DNAs from 12 colon tumor lines, and deletions or rearrangements were seen in 50% (Fig. 2B). Other homozygous deletions (Fig. 1C) were detected in three lung (NCI-H1648, NCI-H460 and NCI-H892) and one breast carcinoma (MDA-231) cell lines. The smallest (~40 kb) was in the cervical carcinoma line CC19. This line was characterized further by DNA sequence analysis and *FHIT* gene expression (see below). Two (HeLa and MDA-231) contained discontinuous deletions, which was surprising since a single deletion should have been sufficient to inactivate a tumor suppressor gene. No deletions were detected in five renal carcinoma lines (KRC/Y, CAKI-1, CAKI-2, ACHN and KV6).

We also observed deletions in non-tumor-derived DNAs. Hybridization with c61G12 to a chromosome 3 hybrid panel revealed a partial deletion in 2A3CT (Fig. 3A). Similarly, a

Figure 1. (**A**) Schematic showing position of the homozygous deletion region with respect to YACs 850A6, 65E7 and 74B2. Positions of *Mlu*I and selected *Xho*I restriction sites are indicated. *Xho*I sites were not determined for YAC 850A6. (**B**) Cosmid and λ clone contig spanning the homozygous deletion region within 3p14. Solid horizontal lines represent cosmid (c) and lambda (λ) inserts, as indicated, along with cleavage sites for *Eco*RI (short vertical bars), *Xho*I (X) and *Sal*I (S) which were mapped within the central 170 kb. The *Mlu*I site corresponding to the site in YAC 74B2 is present in c55D12 and c84F12, although we have not determined its precise location. Positions of selected markers, breakpoints and integration sites for HPV-16 and pSV2neo are indicated across the top. (B) and (C) are drawn to the same scale and positions correspond exactly between the two. (C) Homozygous deletions in tumors and normal genomic DNAs. Horizontal lines denote the extent of deletion in the indicated cell lines. Parallel lines at the ends of each deletion indicate where precise boundaries were not determined. HeLa and MDA231 contain discontinuous deletions indicated by interrupted lines. Exon 5 and the direction of transcription for the *FHIT* gene are indicated at the bottom.


A. YACs
TELOMERE
CENTROMERE
74B2 (450 kb)
65E7 (300 kb)
t(3;8)
850A6 (1330 kb)
homozygous deletion region containing FRA3B
B. Cosmids & Phage
Markers & Rearrangements:
aphidicolin induced break: A5-4
65E7(L)
pSV2neo integration
D3S1300
pSV2neo integration
FHIT-exon 5
distal aphidicolin breakpoint cluster
D3S2723 (R7K145)
HPV integration (3')
HPV integration (5')
189B12(R)
65E7(R)
74B2(R)
EagI
SacII EagI
BssHII
EagI
λ8
λ11
λ104
c31E1
c44A1
c187A2
λ1
λ133
λ55
λ124
λ93
c139C6
c61G12
c210H10
c81E9
λ 58
λ 35
λ 75
c68D2
Sal
c120E3
c64D1
λ 87
c136C3
λ 67
λ 48
λ 34
λ 33
λ 63
λ94
λ38
λ95
λ6
λ37
c55D12
c84F12
λ20
Sequenced Region
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180
C. Deletions
2A3CT
YAC 74B2Δ
CC19
MS751
C33A
SiHa
Caski
HeLa
HeLa
HT29
COLO205
SW480
NCI-H1648
NCI-H460
NCI-H892
MDA231
MDA231
Non-tumor
Cervix
Colon
Lung
Breast
Intron 5
Exon 5
Intron 4
FHIT transcript
3'
5'

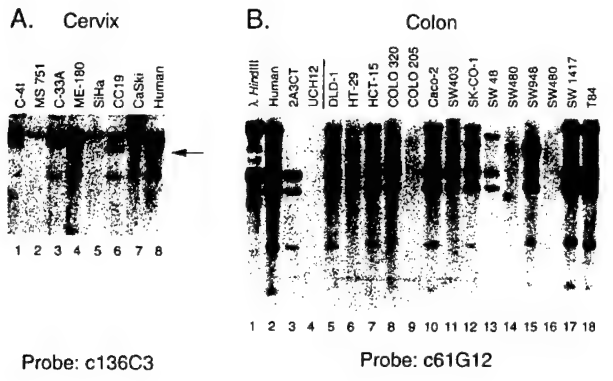


Figure 2. (**A**) Southern analysis of cervical carcinoma cell lines. Lanes 1–7 contain DNA from seven cervical lines, as indicated; lane 8 is a normal human DNA control. DNA samples were digested with *Eco*RI and hybridized with c136C3. Three cell lines were homozygously deleted for c136C3 while the arrow indicates a rearranged band present in CC19. (**B**) Southern analysis of colon carcinoma cell lines. Control lanes (1–4) contain respectively, λ *Hin*dIII size marker, *Hin*dIII-digested human DNA, the chromosome 3-specific hybrids 2A3CT and UCH12. Hybrid 2A3CT contains a single copy of chromosome 3 deleted for all sequences distal to 3p21.3. The chromosome retained in UCH12 is deleted for the entire short (p) arm. Lanes 5–18 contain *Hin*dIII-digested DNAs from 12 colon carcinoma cell lines (note, DLD-1 and HCT-15 are identical). The Southern blot was hybridized with c61G12 which contains *D3S1300* and *FHIT* exon 5. Homozygous deletions were observed in three cell lines; altered bands, which may be indicative of rearrangements, were present in four additional lines.

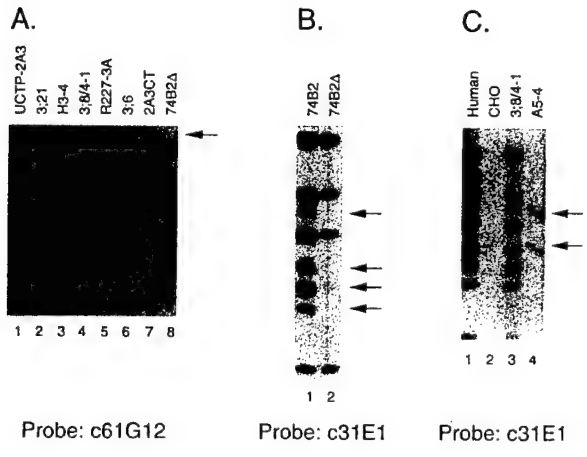


Figure 3. Deletions in genomic DNAs from normal sources. DNA from several hybrids and YAC clones were analyzed with probes from the homozygous deletion region. In (**A**), DNA from seven chromosome 3 hybrids and a deletion variant of YAC clone 74B2 were digested with *Hin*dIII and hybridized with c61G12. The variant band in hybrid H3-4 represents a polymorphism identified in eight of 19 normal DNA samples (not shown). The seven hybrids all retain 3p14.2 by cytogenetic and molecular genetic analyses while missing other specific regions of chromosome 3. (**B**) DNA samples from YACs 74B2 and the deletion variant 74B2D were digested with *Eco*RI and hybridized with c31E1. Four homologous bands present in YAC 74B2 were missing in 74B2D (arrows). In (**C**), *Eco*RI-digested DNA from the hybrid A5-4 (lane 4) was compared with human (lane 1) and hybrid 3;8/4-1 (lane 3) using c31E1.

spontaneous 80 kb deletion in YAC 74B2 spanning the region was identified during single clone purification (74B2D). Figure 3B shows missing bands (arrows) in 74B2D when hybridized with c31E1. This 80 kb deleted segment encompasses the aphidicolin-

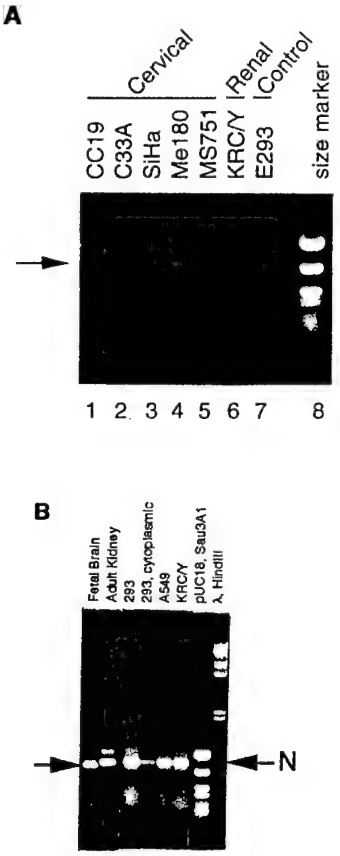


Figure 4. (**A**) RT-PCR analysis of *FHIT* expression in tumor cell lines. Primers designed to amplify the coding region of *FHIT* were used to amplify cDNAs prepared from five cervical carcinoma lines, one renal carcinoma line (KRC/Y) and an E1A-transformed embryonal kidney line (E293), as indicated. Normal products of 638 bp (arrow) were observed in three cell lines, including CC19 with a 40 kb homozygous deletion. A normal sized product was also observed in MS751, although it is amplified weakly. In (**B**), evidence is presented for alternative splicing of *FHIT* mRNA in normal tissues. RT-PCR analysis of *FHIT* expression in fetal brain, adult kidney, E293 cells (whole cells and cytoplasmic fraction), lung tumor A549 and kidney carcinoma cell line KRC/Y were performed as in (A). Both normal sized products (arrow marked N) and smaller products (left arrow) were observed.

induced breakpoint in hybrid A5-4 (Figs 3C and 1C), the ends of several tumor deletions (C33A, SIHA, HeLa, COLO-205 and MDA-231) and the more telomeric of two pSV2neo plasmid insertion sites which preferentially integrated into FRA3B after aphidicolin treatment (20,21).

Correlation of deletions with microsatellite instability

In cervical and colorectal carcinoma cell lines, we observed a trend of inverse correlation between the presence of a 3p14 deletion and reported microsatellite instability (22–26). However, for some lines, we were unable to discern the replication error (RER) status from the literature. Therefore, we subcloned selected lines and tested 10 clones each with up to six microsatellite loci (*D3S1300*, *D3S1210*, *D3S1286*, *D3S1233* and *AFM320yb5*). Instability was accepted if new bands appeared from two or more loci, although there were no cases of only one alteration. Results for samples where we have information on both DNA mismatch repair and 3p14 deletion are shown in

Table 1. While the number of RER+ lines is small, both 3p14 deletions and the RER phenotype are usually discordant. One simple conclusion is that defects associated with microsatellite instability alone are not responsible for the observed deletions. We also observed the highest incidence of deletions in cervical carcinoma lines, where p53 alterations appear very common (27,28). This may relate to the reduced frequency of deletions observed in RER+ colorectal tumors (see Discussion).

Relative position of *FHIT*

To examine the role of the *FHIT* gene in these deletions, we derived primers (FHIT-5′ and FHIT-3′) for a one-stage RT-PCR amplification of the entire coding region (exons 5–9). An expected 638 bp product was amplified after 35 cycles from embryonal kidney 293 cells (Fig. 4A, lane 7). This was used as a probe against the contig which showed that only a single *FHIT* exon, located adjacent to *D3S1300*, was present (Fig. 1B). Comparison of our map with that described by Ohta *et al.* (19) allowed us to conclude that this was exon 5 (also confirmed by DNA sequencing, not shown) and demonstrated that several tumor deletions exclusively affected large introns (Fig. 1C). This occurred in CC19 and SIHA (cervix), NCI-H1648 and NCI-H460 (lung) and MDA-231 (breast).

Analysis of *FHIT* in the tumor lines shown in Figure 4A demonstrated that 3/6, including CC19, contained normal sized bands (arrow). To characterize further the *FHIT* product from CC19, four independent isolates were sequenced. Each contained only normal sequences comprising coding exons 5–9. Thus, the homozygous deletion had no apparent effect on *FHIT* mRNA, suggesting either a different target gene or unselected genomic instability. Additionally, a faint normal sized product was seen in MS751 (Fig. 4A, lane 5), which by Southern blot contains a homozygous deletion including exon 5. A possible explanation is that the deletion is heterogeneous within the cell population suggesting it occurred during culture.

The cervical carcinoma lines C33A and SIHA contained multiple smaller RT-PCR products with no detectable wild-type product. While C33A had an exon 5 deletion that could explain one smaller band (Fig. 1C), the multiplicity of products suggested alternative splicing. The SIHA deletion does not affect an exon although there may be non-recognized discontinuous or overlapping bi-allelic deletions. Despite these obvious differences, the RT-PCR products appeared identical (Fig. 4A, lanes 2 and 3). To examine this, we amplified *FHIT* coding exons from RNA or cDNAs prepared from normal (fetal brain, adult kidney), immortalized (E293) and tumor (A549, KRC/Y) samples (Fig. 4B). While normal *FHIT* (arrow marked N) was observed in each, a smaller product was also seen which predominated in the normal fetal brain cDNA library. When this was cloned and sequenced, exon 8 containing the conserved histidine triad motif and possible zinc-binding site (29) was missing (not shown). This variation had been reported to represent an aberration in squamous cell carcinomas of the head and neck (30). Additional larger products were identified from the adult kidney cDNA library although these have not been characterized. Thus, alternative splicing definitely occurs in normal tissues.

DNA sequence analysis

Since CC19 contained the smallest deletion, we initiated large-scale DNA sequencing studies to identify new genes or structural features that might be responsible for the genomic instability. Six genome equivalents of sequence were obtained from c81E9, c120E3, c136C3 and λ33 and the data assembled. Gaps were closed using directed primers and a 1.3 kb clone gap between c136C3 and λ33 was closed by PCR amplification using λ48 (Fig. 1B) as template. This resulted in 110.4 kb of contiguous sequence (GenBank U66722) at an accuracy of ~99.6%. Predicted restriction maps from the assembled sequence also matched those determined experimentally. Sequence analysis (shown schematically in Fig. 5) included similarity searches, repetitive sequence identification and exon prediction.

Gene search

Overall, the sequence is AT-rich (61.1%) and very depleted in CG dinucleotides. No identities to known genes were seen. GRAIL2 predicted 15 exons of which six had moderate (<0.4), five had good (0.4–0.6) and four had excellent (>0.6) scores. Four putative exons were coincident with repeat sequences and two others (positions 8.0 and 71.5 kb) were coincident with Genemark predicted coding segments. Genemark (31) was utilized in order to implement a matrix for higher A–T content regions. While 61 potential coding segments were identified, no significant similarities were observed. Many predicted coding segments clearly occurred within LINE and MER elements, and others not directly within repeats nevertheless demonstrated similarity to LINEs. Four putative exons were clustered near position 107 kb, one of which showed perfect identity with a partially sequenced cDNA, EST-N70372. However, Northern analysis failed to identify a transcript from this region (data not shown) and the cDNA sequence included a portion of the LINE element at position 108 kb. Based on end sequences from the cDNA clone and insert length, the cDNA was co-linear with genomic DNA. Further, the 3′ end of the cDNA was coincident with a poly(A) tract in genomic DNA, all suggesting this represented an unprocessed transcript, or more likely resulted from false priming of contaminating DNA. BLASTN searches determined two additional regions with similarity to non-annotated cDNAs (positions 58.2 and 64.3 kb). However, neither showed similarity to known genes nor were directly superimposed on predicted exons. That these sequences were observed adjacent to, rather than superimposed on, predicted exons could be due to conservative prediction algorithms used which underestimate the extent of many exons. Alternatively, the homologies could be due to infrequent repeats. We note the presence of five remaining putative exons, denoted in Figure 5, having high probability scores and which do not overlap repeats.

Repeat sequence analysis

Analysis with Pythia (32) showed that 20.2% of the sequenced region is comprised of known repeats, a level comparable with other regions chosen for comparison (Table 2). However, in this 110.4 kb, LINE and MER elements make up the bulk of repeats. Intriguingly, the repeat composition is very similar to a 152 kb sequence from the Fragile-X region (33) and differs from arbitrarily selected segments in 3p21.3 and 4p16.3 (Table 2). In particular, λ33 (from position 96 to 110.4 kb) within the CC19 deletion is nearly identical in LINE element composition to the Fragile-X region. Similarly, both Fragile-X and FRA3B have a low level of predicted coding regions. Speculatively, these similarities may influence the observed tendency of both regions to undergo breakage. However, there are obvious differences between the two sites, notably the presence of a triplet repeat and

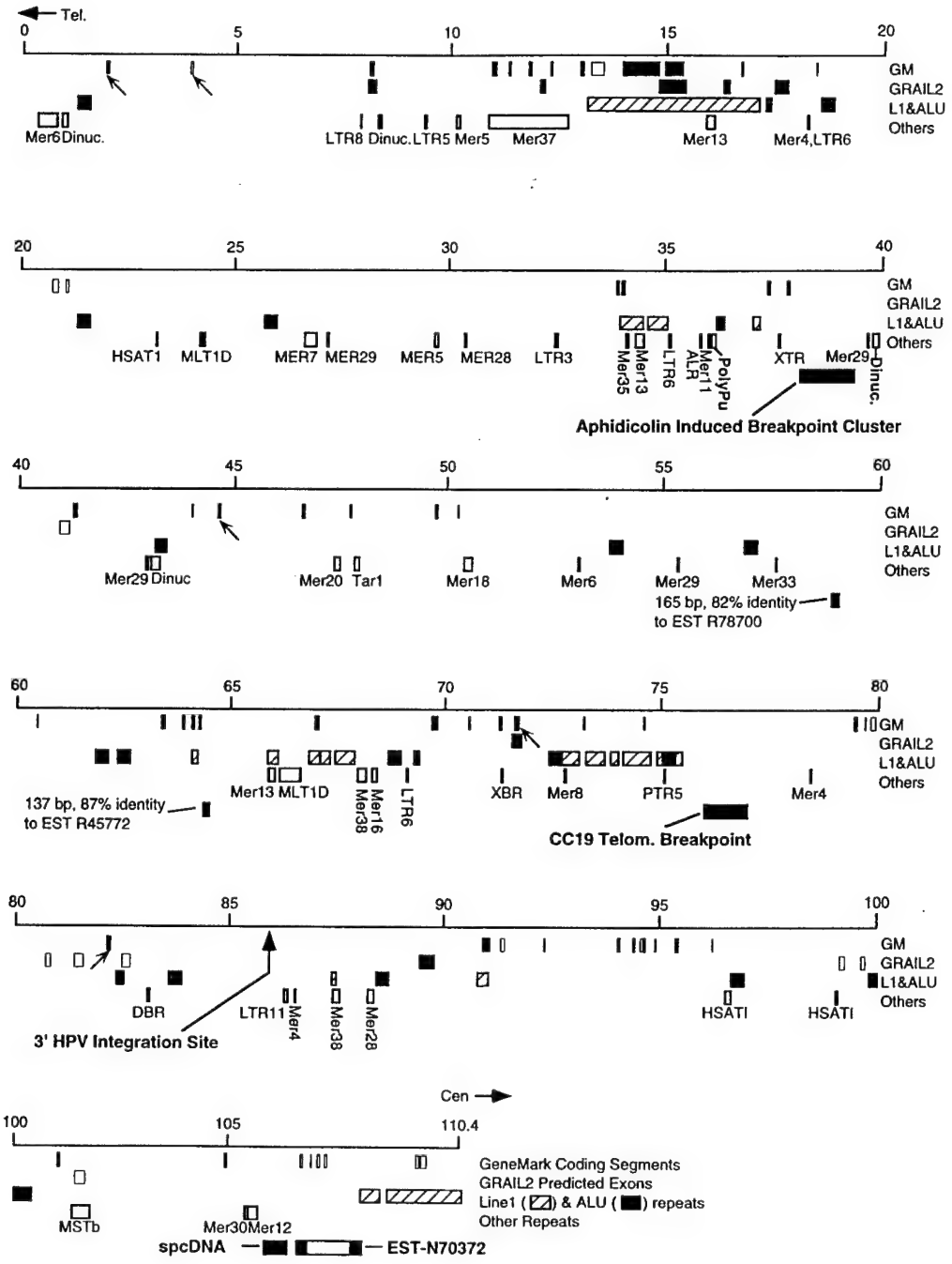


Figure 5. Assembled sequence data totaling 110 435 bp. GeneMark and GRAIL2 predicted exons are indicated (filled boxes = direct strand; open boxes = complementary strand). Repetitive sequences homologous to LINE (L1) and Alu elements are designated by boxes filled with diagonal lines and shading, respectively. Other repetitive elements are labeled. These largely consist of MER elements (medium reiterated repeats) and LTRs (long terminal repeats). Information about particular repetitive elements can be found in Repbase (ftp://ncbi.nlm.nih.gov/repository/repbase). Unique features of the sequence are indicated on the lowest row.

methylated CpG island in Fragile-X (34) and the common versus rare nature of FRA3B.

Other features

Position 38 kb contains a cluster of terminal deletion breakpoints induced by aphidicolin treatment (GenBank U46001) further linking our sequence to FRA3B. Flanking these aphidicolin breakpoints is a long polypurine tract and two extended variable dinucleotide repeats. The telomeric breakpoint of the CC19 deletion lies within a 1.0 kb region at position 76 kb between L1 and MER/LTR elements. Within the CC19 deletion region, near position 86 kb, is an HPV16 integration site from a cervical carcinoma (35). It is of note that this integration was associated with an interstitial deletion (35,36) which we now know would not affect *FHIT* coding sequences. Interestingly, a very signifi-

cant similarity ($1e^{-77}$) to a small polydispersed circular (spc) DNA was observed at 106 kb (GenBank X96885). Figure 6 shows a FASTA alignment over 585 bp. The region is 72.1% identical and contains interspersed blocks having up to 89% identity (e.g., spcDNA bp 121–207). This spcDNA, which appears non-repetitive, was isolated from a tuberous sclerosis-associated angiofibroma (I. Hinkel-Schreiner, Ph.D. Thesis). Characteristics of spcDNAs include derivation from chromosomal sequences (37), association with clustered repeats [such as β-satellites and other clustered elements although a single family member may be predominantly involved (38)] and elevation in conditions associated with genomic instability such as Fanconi's anemia (39). spcDNAs are also increased by DNA-damaging agents (37) and inhibitors of DNA and protein synthesis including the fragile site inducer aphidicolin (40). Given the limited number of spcDNAs that have been sequenced, this similarity may be biologically important.

DISCUSSION

Using a ~300 kb cosmid/λ contig, located ~150 kb telomeric to the 3;8 translocation breakpoint, we have identified homozygous deletions in various carcinoma cell lines that overlap the most inducible common fragile site in the genome, FRA3B. From various aphidicolin-induced breakpoints and plasmid or viral integration sites (Fig. 1), FRA3B represents a region rather than a single site. Studies by Wilke *et al.* (36) and Smith *et al.* (41) indicate that some clustering of breakpoints may occur. However, their rearrangements were induced by aphidicolin in a single chromosome 3-containing hybrid and, unlike the interstitial deletions we observed in tumor and non-tumor samples, appear to represent terminal breaks. Where we have defined the boundaries for the carcinoma-associated deletions accurately, one or both are contained within the FRA3B region.

Our initial hypothesis was that the smallest deletion (CC19) would contain elements of a tumor suppressor gene. During our studies, Ohta *et al.* (19) identified the *FHIT* gene with reported abnormalities in RT-PCR products. However, *FHIT* has similarity to a yeast di-adenosine hydrolase which would represent an unexpected function for a tumor suppressor gene. Our results indicate that *FHIT* is not the target of these deletions. First, the CC19 deletion does not involve *FHIT* coding sequences (Fig. 1B and C) and, based on RT-PCR and cDNA sequence analysis, the coding portion of the *FHIT* transcript is normal. Second, we have observed deletions in genomic DNAs from non-tumor sources (Fig. 3A). The somatic cell hybrid 2A3CT was formed by spontaneous terminal deletion at 3p21.3 from a non-tumor-derived chromosome 3 hybrid, UCTP2A3 (11). However, 2A3CT also acquired a 3p14 interstitial deletion overlapping an aphidicolin-induced breakpoint, a pSV2neo integration into FRA3B (20) and the telomeric borders of several carcinoma-associated deletions. We also identified an overlapping deletion in an unselected subclone from YAC 74B2. While neither the hybrid nor YAC 3p14 DNA segments are in a 'normal' background, they clearly are unselected from a tumorigenic standpoint. We note that 'hotspots' of recombination in human DNA can be maintained in a yeast background (42), thus it is not unreasonable that unstable regions may behave similarly or more so. Third, discontinuous deletions appear common in this region (i.e. HeLa, MDA231), both from our analysis and from that reported by Ohta *et al.* (19). Multiple deletions might be expected if there were no common target gene and if the region was unstable. Fourth, we have observed that *FHIT* undergoes alternative splicing in normal tissues (Fig. 4B) which explains some previously reported abnormal PCR products (19,30,43). Thiagalingam *et al.* (44) recently reported lack of *FHIT* involvement in colorectal carcinomas and suggested that PCR artifacts might be responsible for some previously observed alterations (19,43). It also seems likely that many deletions would have been missed in their study since only a few markers were tested. One of the possible features that suggested *FHIT* could be a tumor suppressor gene was that it crossed the hereditary renal carcinoma-associated 3;8 breakpoint (19). However, we found no alterations in RT-PCR products from five renal carcinoma cell lines. Moreover, Bugert *et al.* (manuscript submitted) have observed normal *FHIT* transcripts and no point mutations in a large series of renal cancers. Thus, *FHIT* does not appear to be involved in renal carcinoma. With respect to other possible target genes, from our sequencing studies we identified a 100% identity to two expressed sequence tags from a liver/spleen library. However, we determined that both clones were identical, were not expressed using a commercial Northern blot (Clontech), were co-linear with genomic DNA containing a poly(A) tract corresponding to their 3′ end, encoded no significant open reading frame and overlapped a partial LINE element.

While other tumor suppressor genes may exist within FRA3B, an alternative possibility is that the deletions are due to primary genomic instability affecting a particularly susceptible region. This hypothesis is consistent with several of the observations reported here including their discontinuity and occurrence in non-tumor cell lines. By using numerous probes, we were able to identify a high frequency of homozygous deletions, especially in cervical carcinomas where p53 inactivation is very common (18). In this regard, it is interesting that we observed an inverse correlation between 3p14 deletions and microsatellite instability (RER+). Importantly, p53 mutations, which have been shown to destabilize the genome (45,46), appear infrequent in RER+ colorectal carcinomas and gastric tumors [(47,48) and P. Cottu, presented at Cancer and the Cell Cycle, Lausanne, Switzerland 1996]. Thus, these findings would be consistent, at least in part, with 3p14 deletions resulting from the genomic instability which accompanies p53 inactivation. We note that although CC19 is HPV negative, it expresses a mutant p53 protein (53).

What have we learned from the DNA sequence analysis to date? First, is that the region is high in A–T content with frequent LINEs and MER repeats, and is conversely low in Alu sequences and confirmed genes. In contrast to the reported rare folate-sensitive sites which are associated with expanded CGG repeats (49,50), FRA3B does not contain an expanded triplet repeat or methylated CpG island; nor did we identify any telomeric repeats which have been suggested as a possible cause of breaks (51). However, we do note that over the 110.4 kb region there is an overall repeat sequence similarity to 152 kb from Fragile-X. *In vitro*, expanded CGG repeats have been shown to inhibit DNA replication (52). Whether or not specific sequences within FRA3B may have a similar effect on replication awaits experimental testing. Our discovery of a strong spcDNA homology in FRA3B may therefore not be coincidental. While we do not know if the spcDNA site is a primary cause of FRA3B instability or simply a marker for this property, the DNA sequences reported should provide the means to test this.

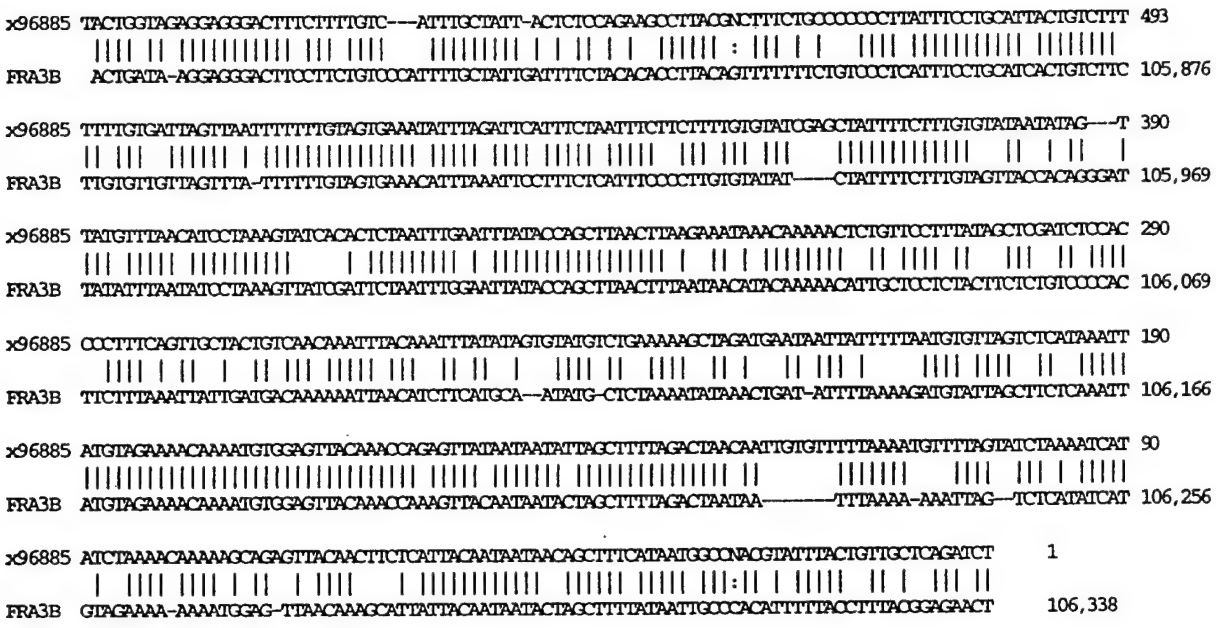


Figure 6. FASTA alignment of 585 bp of spcDNA clone (X96885) and FRA3B.

Table 2. Summary of sequence features

	Sequence length	GC content	Total repeats	Alu	LINE	MER	Predicted exons	Confirmed genes
FRA3B (110.4 kb sequence)	110 434	38.9%	20.2%	4.9%	9.0%	3.9%	3.4%	0.0%
FRA3B (λ33 only)	13 121	37.0%	26.6%	3.5%	17.5%	2.1%	2.3%	0.0%
Xq27.3 (Fragile X)	152 351	39.0%	26.3%	5.9%	16.3%	3.0%	3.4%	2.8%
3p21.3 (HLuca_14)	36 597	58.0%	19.4%	18.1%	0.0%	0.4%	9.5%	12.9%
4p16.3 (Huntington's)	32 100	54.9%	16.1%	10.8%	3.5%	1.4%	5.6%	7.6%

The sequences used for comparison were derived from Xq27.3 (contains the Fragile-X region, GenBank L29074) and 4p16.3 (in the region of Huntington's disease, GenBank Z69837). The sequence from 3p21.3 was downloaded from ftp://genome.wustl.edu/pub/gsc1/sequence/st.louis/human/shotgun/3/H_LUCA14.seq.

MATERIALS AND METHODS

Nucleic acid sources and manipulations

Cell lines. Breast, colon and cervical carcinoma lines were obtained from the American Type Culture Collection. The cervical carcinoma cell line, CC19, was established as described (53). Lung tumor lines were obtained through the Colorado Lung Cancer SPORE Tissue Culture Core Laboratory. Normal cell lines included the human lymphoblastoid cell lines TL8229 and AG4103 and the E1A-immortalized human embryonic kidney cell line, E293. The somatic cell hybrids have been described previously (54).

Libraries. The gridded chromosome 3-specific cosmid library (55) was obtained from Lawrence Livermore National Laboratories. The 850A6 YAC subclone library was previously described (16). Fetal brain and adult kidney cDNA libraries were obtained from Clontech.

DNA and RNA isolations. DNA was isolated from cell lines by standard methods. Cosmid DNAs were isolated using alkaline lysis; preparations used for DNA sequence analysis were purified further by CsCl gradient centrifugation. DNA was isolated from single or pooled phage clones by the Grossberger method (56). RNA was isolated from cell lines when the cultures reached 90% of confluency using the RNA-STAT-60 kit from Tel-Test, Inc. (Friendswood, TX).

Library screening, contig assembly and hybridization analysis. The gridded cosmid library was spotted onto filters at high density (1536 clones per filter) and hybridized using standard techniques with a 370 kb *Mlu*I fragment derived from YAC 74B2. The 850A6 YAC phage library (16) was screened using the same probe. Resulting clones were assembled into a contiguous segment by hybridization analysis using end probes and total inserts followed by analysis with the software tool SEGMAP. Complete and partial digestion analyses (57) were used to restriction map the central 170 kb of the contig prior to DNA sequence analysis.

Replication error (RER) analysis

Biotinylated primers were obtained from Research Genetics, Huntsville, AL. PCR amplifications were performed with 40 ng of template DNA utilizing hot start and touch down procedures. After separation on denaturing polyacrylamide gels, PCR products were detected using the New England BioLabs Phototope™ Detection Kit. Alleles were scored by visual inspection of band patterns.

RT-PCR analysis of *FHIT* gene expression

RT-PCR was performed using primers FHIT-5′ (5′-CTCGAATTCTTAGACCCTCTATAAAAGC-3′) and FHIT-3′ (5′-CTGATTCAGTTCCTCTTG-3′) derived from non-coding exons 4 and 10, respectively. First strand synthesis was accomplished with 1–3 mg of total RNA and the Superscript II kit (Life Technologies). Subsequent amplification utilized one-fifth of the reverse transcriptase reaction together with the *FHIT* primers. Standard PCR conditions of 94°C denaturation (1 min), 55°C annealing (1 min) and 72°C elongation for 35 cycles were employed. PCR products were subcloned into the *Eco*RI site of pBluescript II SK$^+$ using an introduced *Eco*RI site present in the FHIT-5′ primer and a natural site located 21 bp downstream of the FHIT stop codon.

DNA sequence determination

Cosmid and phage clones were sequenced using a random shotgun subcloning and end-sequencing strategy. Clone DNA was sonicated and size selected by LMP-agarose gel electrophoresis. Recovered fragments of 1–2 kb were end-repaired with Klenow fragment of *Escherichia coli* DNA polymerase I and T4 DNA polymerase, ligated into the phosphatased *Eco*RV site of pBluescript II and transformed into *E.coli* DH10B. Amp^R/β-Gal$^-$ subclones were grown in 3 ml of TB for isolation of sequencing templates. To eliminate vector sequences from the subclone library, phage inserts were amplified by long-range PCR. Inserts were gel purified, ^{32}P labeled and hybridized to subclone libraries. Positive subclones were sequenced using an ABI 373 or 377 and the ABI Prism dye terminator cycle sequencing kit. Chromatograms were transferred to a SUN workstation, analyzed and assembled using PHRED and PHRAP (from Dr Phil Green). Gaps were closed by primer walking. Based on independently obtained overlapping contigs, as well as analysis of cosmid vector sequences in some subclones, sequencing accuracy was at least 99.6%.

Sequence analysis

Assembled sequences were analyzed for database similarities by BLASTN and BLASTX, searching the nr and dbest databases with default parameters. Strongly similar sequences ($P < e^{-30}$) were retrieved from GenBank for further analysis. Homologies to repeat sequences were found using Pythia (32). Potential exons were identified using GRAIL2 (58) and GeneMark (31) programs. GeneMark was run using 5th order matrices trained on sequences with GC content similar to the sequences checked. Sequence alignments were prepared using the FASTA program from the Wisconsin Package (GCG).

ACKNOWLEDGEMENTS

This work was supported by DAMD17-94-J-4391 (HAD, RMG), CA51395 (HPK, MRM) and the DNA sequencing studies were supported by HG000358 (H.A.D., R.M.G.). We gratefully acknowledge support from the Colorado Cancer League (2531718) to F.B. Additional Core support was provided through the University of Colorado Cancer Center (CA 46934) and by Dr Wilbur Franklin through the Lung Cancer Spore.

REFERENCES

1. Ogawa, O., Kakehi, Y., Ogawa, K., Koshiba, M., Sugiyama, T. and Yoshida, O. (1991) Allelic loss at chromosome 3p characterizes clear cell phenotype of renal cell carcinoma. *Cancer. Res.*, **51**, 949–953.
2. Devilee, P., van den Broek, M., Kuipers-Dijkshoorn, N., Kolluri, R., Khan, P.M., Pearson, P.L. and Cornelisse, C.J. (1989) At least four different chromosomal regions are involved in loss of heterozygosity in human breast carcinoma. *Genomics*, **5**, 554–560.
3. Hibi, K., Takahashi, T., Yamakawa, K., Ueda, R., Sekido, Y., Ariyoshi, Y., Suyama, M., Takagi, H. and Nakamura, Y. (1992) Three distinct regions involved in 3p deletion in human lung cancer. *Oncogene*, **7**, 445–449.
4. Lothe, R.A., Fossa, S.D., Stenwig, A.E., Nakamura, Y., White, R., Borresen, A.L. and Brogger, A (1989) Loss of 3p or 11p alleles is associated with testicular cancer tumors. *Genomics*, **5**, 134–138.
5. Yang-Feng, T.L., Han, H., Chen, K.C., Li, S.B., Claus, E.B., Carcangiu, M.L., Chambers, S.K., Chambers, J.T. and Schwartz, P.E. (1993) Allelic loss in ovarian cancer. *Int. J. Cancer*, **54**, 546–551.
6. Kohno, T., Takayama, H., Hamaguchi, M., Takano, H., Yamaguchi, N., Tsuda, H., Hirohashi, S., Vissing, H., Shimizu, M., Oshimura, M. *et al.* (1993) Deletion mapping of chromosome 3p in human uterine cervical cancer. *Oncogene*, **8**, 1825–1832
7. Maestro, R., Gasparotto, D., Vukosavljevic, T., Barzan, L., Sulfaro, S. and Boiocchi, M. (1993) Three discrete regions of deletion at 3p in head and neck cancers. *Cancer Res.*, **53**, 5775–5779.
8. Yamakawa, K., Takahashi, T., Horio, Y., Murata, Y., Takahashi, E., Hibi, K., Yokoyama, S., Ueda, R. and Nakamura, Y. (1993) Frequent homozygous deletions in lung cancer cell lines detected by a DNA marker located at 3p21.3-p22. *Oncogene*, **8**, 327–330.
9. Drabkin, H.A., Mendez, M.J., Rabbitts, P.H., Varkony, T., Bergh, J., Schlessinger, J., Erickson, P. and Gemmill, R.M. (1992) Characterization of the submicroscopic deletion in the small- cell lung carcinoma (SCLC) cell line U2020. *Genes, Chromosomes Cancer*, **5**, 67–74.
10. Rabbitts, P., Bergh, J., Douglas, J., Collins, F. and Waters, J. (1990) A submicroscopic homozygous deletion at the D3S3 locus in a cell line isolated from a small cell lung carcinoma.. *Genes, Chromosomes Cancer*, **2**, 231–238.
11. Roche, J., Boldog, F., Robinson, M., Robinson, L., Varella-Garcia, M., Swanton, M., Waggoner, B., Fishel, R., Franklin, W., Gemmill, R. and Drabkin, H. (1996) Distinct 3p21.3 deletions in lung cancer and identification of a new human semaphorin. *Oncogene*, **12**, 1289–1297.
12. Lisitsyn, N.A., Lisitsina, N.M., Dalbagni, G., Barker, P., Sanchez, C.A., Gnarra, J., Linehan, W.M., Reid, B.J. and Wigler, M.H. (1995) Comparative genomic analysis of tumors: detection of DNA losses and amplification. *Proc. Natl Acad. Sci. USA*, **92**, 151–155.
13. Yokota, J., Tsukada, Y., Nakajima, T., Gotoh, M., Shimosato, Y., Mori, N., Tsunokawa, Y., Sugimura, T. and Terada, M. (1989) Loss of heterozygosity on the short arm of chromosome 3 in carcinoma of the uterine cervix. *Cancer Res.*, **49**, 3598–3601.
14. Cohen, A.J., Li, F.P., Berg, S., Marchetto, D.J., Tsai, S., Jacobs, S.C. and Brown, R.S. (1979) Hereditary renal-cell carcinoma associated with a chromosomal translocation. *N. Engl. J. Med.*, **301**, 592–595.

15. Glover, T.W., Berger, C., Coyle, J. and Echo, B. (1984) DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Hum. Genet.*, **67**, 136–142.
16. Boldog, F.L., Gemmill, R.M., Wilke, C.M., Glover, T.W., Nilsson, A.S., Chandrasekharappa, S.C., Brown, R.S., Li, F.P. and Drabkin, H.A. (1993) Positional cloning of the hereditary renal carcinoma 3;8 chromosome translocation breakpoint. *Proc. Natl Acad. Sci. USA*, **90**, 8509–8513.
17. Wilke, C.M., Guo, S.W., Hall, B.K., Boldog, F., Gemmill, R.M., Chandrasekharappa, S.C., Barcroft, C.L., Drabkin, H.A. and Glover, T.W. (1994) Multicolor FISH mapping of YAC clones in 3p14 and identification of a YAC spanning both FRA3B and the t(3;8) associated with hereditary renal cell carcinoma. *Genomics*, **22**, 319–326.
18. Milde-Langosch, K., Albrecht, K., Joram, S., Schlechte, H., Giessing, M. and Loning, T. (1995) Presence and persistence of HPV infection and p53 mutation in cancer of the cervix uteri and the vulva. *Int. J. Cancer*, **63**, 639–645.
19. Ohta, M., Inoue, H., Cotticelli, M.G., Kastury, K., Baffa, R., Palazzo, J., Siprashvili, Z., Mori, M., McCue, P., Druck, T., Croce, C.M. and Huebner, K. (1996) The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. *Cell*, **84**, 587–597.
20. Rassool, R.V., Le Beau, M.M., Shen, M.-L., Neilly, M.E., Espinosa, R, III, Ong, S.T., Boldog, F., Drabkin, H., McCarroll, R. and McKeithan, T.W. (1996) Direct cloning of DNA sequences from the common fragile site region at chromosome band 3p14.2. *Genomics*, in press,
21. Rassool, F.V., McKeithan, T.W., Neilly, M.E., van Melle, E., Espinosa, R. and LeBeau, M.M. (1991) Preferential integration of marker DNA into the chromosomal fragile site at 3p14: an approach to cloning fragile sites. *Proc. Natl Acad. Sci. USA*, **88**, 6657–6661.
22. Bhattacharyya, N.P., Ganesh, A., Phear, G., Richards, B., Skandalis, A. and Meuth, M. (1996) Molecular analysis of mutations in mutator colorectal carcinoma cell lines. *Hum. Mol. Genet.*, **4**, 2057–2064.
23. Boyer, J.C., Umar, A., Risinger, J.I., Lipford, J.R., Kane, M., Yin, S., BarrettJ.C., Kolodner, R.D. and Kunkel, T.A. (1995) Microsatellite instability, mismatch repair deficiency, and genetic defects in human cancer cell lines. *Cancer Res.*, **55**, 6063–6070.
24. da Costa, L.T., Liu, B., el-Deiry, W., Hamilton, S.R., Kinzler, K.W., VogelsteinB., Markowitz, S., Willson, J.K., de la Chapelle, A., Downey, K.M. *et al.* (1995) Polymerase delta variants in RER colorectal tumours [letter]. *Nature Genet.*, **9**, 10–11.
25. Umar, A., Boyer, J.C., Thomas, D.C., Nguyen, D.C., Risinger, J.I., Boyd, J., Ionov, Y., Perucho, M. and Kunkel, T.A. (1994) Defective mismatch repair in extracts of colorectal and endometrial cancer cell lines exhibiting microsatellite instability. *J. Biol. Chem.*, **269**, 14367–14370.
26. Larson, A.A., Kern, S., Sommers, R.L., Yokota, J., Cavenee, W.K. and Hampton, G.M. (1996) Analysis of replication error (RER+) phenotypes in cervical carcinoma. *Cancer Res.*, **56**, 1426–1431.
27. Johnson, E.S., Ma, P.C., Ota, I.M. and Varshavsky, A. (1995) A proteolytic pathway that recognizes ubiquitin as a degradation signal. *J. Biol. Chem.*, **270**, 17442–17456.
28. Band, V., Dalal, S., Delmolino, L. and Androphy, E.J. (1993) Enhanced degradation of p53 protein in HPV-6 and BPV-1 E6-immortalized human mammary epithelial cells. *EMBO J.*, **12**, 1847–1852.
29. Lima, C.D., Klein, M.G., Weinstien, I.B. and Hendrickson, W.A. (1996) Three-dimensional structure of human protein kinase C interacting protein 1, a member of the HIT family of proteins. *Proc. Natl Acad. Sci. USA*, **93**, 5357–5362.
30. Virgilio, L., Shuster, M., Gollin, S.M., Veronese, M.L., Ohta, M., Huebner, K. and Croce, C.M. (1996) FHIT gene alterations in head and neck squamous cell carcinomas. *Proc. Natl Acad. Sci. USA*, **93**, 9770–9775.
31. Borodovsky, M. and McIninch, J. (1993) GenMark: parallel gene recognition for both DNA strands. *Comput. Chem.*, **17**, 123–133.
32. Milosavljevic, A. (1994) In *Imperial Cancer Research Fund Handbook of Genome Analysis*. Vol. 1, Blackwell Scientific Publications.
33. Verkerk, A.J., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P., Pizzuti, A., Richards, S., Victoria, M.F., Zhang, F.P. *et al.* (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*, **65**, 905–914.
34. Kremer, E.J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S. T., Sutherland, G.R. and Richards, R.I. (1991) Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science*, **252**, 1711–1714.
35. Wagatsuma, M., Hashimoto, K. and Matsukura, T. (1990) Analysis of integrated human papillomavirus type 16 DNA in cervical cancers: amplification of viral sequences together with cellular flanking sequences. *J. Virol.*, **64**, 813–821.
36. Wilke, C.M., Hall, B.K., Hoge, A., Paradee, W., Smith, D.I. and Glover, T.W. (1996) FRA3B extends over a broad region and contains a spontaneous HPV16 integration site: direct evidence for the coincidence of viral integration sites and fragile sites. *Hum. Mol. Genet.*, **5**, 187–195.
37. Cohen, S. and Lavi, S. (1996) Induction of circles of heterogeneous sizes in carcinogen-treated cells: two-dimensional gel analysis of circular DNA molecules. *Mol. Cell. Biol.*, **16**, 2002–2014.
38. Misra, R., Shih, A., Rush, M., Wong, E. and Schmid, C.W. (1987) Cloned extrachromosomal circular DNA copies of the human transposable element THE-1 are related predominantly to a single type of family member. *J. Mol. Biol.*, **196**, 233–243.
39. Motejlek, K., Schindler, D., Assum, G. and Krone, W. (1993) Increased amount and contour length distribution of small polydisperse circular DNA (spcDNA) in Fanconi anemia. *Mutat. Res.*, 293, 205–214.
40. Sunnerhagen, P., Sjoberg, R.M. and Bjursell, G. (1989) Increase of extrachromosomal circular DNA in mouse 3T6 cells on perturbation of DNA synthesis: implications for gene amplification. *Somat. Cell. Mol. Genet.*, **15**, 61–70.
41. Paradee, W., Mullins, C., He, Z., Glover, T., Wilke, C., Opalka, B. and Schutte, J. (1995) Precise localization of aphidicolin-induced breakpoints on the short arm of human chromosome 3. *Genomics*, **27**, 358–361.
42. Klein, S., Zenvirth, D., Sherman, A., Ried, K., Rappold, G. and Simchen, G. (1996) Double-strand breaks on YACs during yeast meiosis may reflect meiotic recombination in the human genome. *Nature Genet.*, **13**, 481–484.
43. Sozzi, G., Veronese, M.L., Negrini, M., Baffa, R., Cotticelli, M.G., Inoue, H., Pilotti, S., De Gregorio, L., Pastorino, U., Pierotti, M. A., OhtaM., Huebner, K. and Croce, C.M. (1996) The FHIT gene 3p14.2 is abnormal in lung cancer. *Cell*, **85**, 17–26.
44. Thiagalingam, S., Lisitsyn, N.A., Hamaguchi, M., Wigler, M.H., Willson, J.K.V., Markowitz, S.D., Leach, F.S., Kinzler, K.W. and Vogelstein, B. (1996) Evaluation of the FHIT gene in colorectal cancers. *Cancer Res.*, **56**, 2936–2939.
45. Yin, Y., Tainsky, M.A., Bischoff, F.Z., Strong, L.C. and Wahl, G.M. (1992) Wild-type p53 restores cell cycle control and inhibits gene amplification in cells with mutant p53 alleles. *Cell*, **70**, 937–948.
46. Livingstone, L.R., White, A., Sprouse, J., Livanos, E., Jacks, T. and Tlsty, T.D. (1992) Altered cell cycle arrest and gene amplification potential accompany loss of wild-type p53. *Cell*, **70**, 923–935.
47. Renault, B., Calistri, D., Buonsanti, G., Nanni, O., Amadori, D. and Ranzani, G.N. (1996) Microsatellite instability and mutations of p53 and TGF-beta RII genes in gastric cancer. *Hum. Genet.*, in press.
48. Wu, C., Akiyama, Y., Imai, K., Miyake, S., Nagasaki, H., Oto, M., Okabe, S., Iwama, T., Mitamura, K. and Masumitsu, H. (1994) DNA alterations in cells from hereditary non-polyposis colorectal cancer patients. *Oncogene*, **9**, 991–994.
49. Pintado, E., de Diego, Y., Hmadcha, A., Carrasco, M., Sierra, J. and Lucas, M. (1995) Instability of the CGG repeat at the FRAXA locus and variable phenotypic expression in a large fragile X pedigree. *J Med. Genet*, **32**, 907–908.
50. Nancarrow, J.K., Holman, K., Mangelsdorf, M., Hori, T., Denton, M., Sutherland, G.R. and Richards, R.I. (1995) Molecular basis of p(CCG)n repeat instability at the FRA16A fragile site locus. *Hum. Mol. Genet.*, **4**, 367–372.
51. Bouffler, S., Silver, A., Papworth, D., Coates, J. and Cox, R. (1993) Murine radiation myeloid leukaemogenesis: relationship between interstitial telomere-like sequences and chromosome 2 fragile sites. *Genes, Chromosomes Cancer*, **6**, 98–106.
52. Usdin, K. and Woodford, K.J. (1995) CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis *in vitro*. *Nucleic Acids Res.*, **23**, 4202–4209.
53. Mullokandov, M.R., Kholodilov, N.G., Atkin, N.B., Burk, R.D., Johnson, A.B. and Klinger, H.P. (1996) Genomic alterations in cervical carcinoma: losses of chromosome heterozygosity and human papilloma virus tumor status. *Cancer Res.*, **56**, 197–205.
54. Drabkin, H.A., Wright, M., Jonsen, M., Varkony, T., Jones, C., Sage, M., Gold, S. and Erickson, P. (1990) Development of a somatic cell hybrid mapping panel and molecular probes for human chromosome 3. *Genomics*, **8**, 435–446.

55. Ioannou, P.A., Amemiya, C.T., Garnes, J., Kroisel, P.M., Shizuya, H., Chen, C., Batzer, M.A. and de Jong, P.J. (1994) A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nature Genet.*, **6**, 84–89.
56. Grossberger, D. (1987) Minipreps of DNA from bacteriophage lambda. *Nucleic Acids Res.*, **15**, 6737.
57. Sambrook, J. Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
58. Xu, Y., Einstein, J.R., Mural, R.J., Shah, M. and Uberbacher, E.C. (1994) *Proceedings of the 2nd International Conference on Intelligent Systems for Molecular Biology*. Vol. 1, AAAI Press, pp. 376–384.

NOTE ADDED IN PROOF

Sequencing of the CC19 deletion region has been completed with the finalization of lambda 94, a phage proximal to lambda 33. This 18 810 bp sequence maintains the same characteristics as the balance of FRA3B. It is 63% AT and 27% is comprised of known repeats with the L1 type elements predominating.

GENES, CHROMOSOMES & CANCER 21:281–289 (1998)

PERSPECTIVE ARTICLE

An *FHIT* Tumor Suppressor Gene?

Michelle M. Le Beau,[1]* Harry Drabkin,[2] Thomas W. Glover,[3] Robert Gemmill,[2] Feyruz V. Rassool,[1] Timothy W. McKeithan,[4] and David I. Smith[5]

[1]Section of Hematology/Oncology, University of Chicago, Chicago, Illinois
[2]Division of Medical Oncology, University of Colorado Health Sciences Center, Denver, Colorado
[3]Departments of Pediatrics and Human Genetics, University of Michigan, Ann Arbor, Michigan
[4]Departments of Pathology, and Radiation and Cellular Oncology, University of Chicago, Chicago, Illinois
[5]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota

The FRA3B at 3p14.2 is the most common of the constitutive aphidicolin-inducible fragile sites. Using independent approaches, four groups of investigators have cloned and characterized this fragile site. The results of these studies have revealed that the FRA3B differs from other heretofore cloned rare fragile sites. First, instability as manifested by chromosome breakage occurs over a large region of DNA, encompassing at least 500 kb. Second, sequence analysis has not revealed trinucleotide repeat motifs, characteristic of the rare fragile sites. In addition to containing the FRA3B, band 3p14 is also likely to contain a tumor suppressor gene, as evidenced by the presence of deletions, rearrangements, and allele loss in a variety of human tumors, including lung, renal, nasopharyngeal, cervical, and breast carcinomas. The recently cloned *FHIT* gene in 3p14.2 is a promising candidate tumor suppressor gene, since aberrant *FHIT* transcripts have been found in a significant proportion of cancer-derived cell lines and primary tumors of the digestive and respiratory tracts. Nonetheless, several lines of evidence garnered over the past year have called into question the role of *FHIT* as a classical tumor suppressor gene, and raised the question of whether its apparent involvement simply reflects its location within an unstable region of the genome. In the following study, we have summarized the evidence in support of *FHIT* as a tumor suppressor gene as well as evidence against such a role, and the experimental evidence needed to demonstrate that *FHIT* functions as a tumor suppressor gene in the pathogenesis of human tumors. The paradigm of *FHIT* emphasizes that confirming the role of a candidate tumor suppressor gene may prove difficult, particularly for those genes that are located in genetically unstable regions. *Genes Chromosomes Cancer 21:281–289, 1998.*

Chromosomal Fragile Sites

Chromosomal fragile sites are specific loci that are especially susceptible to forming gaps, breaks, or rearrangements in metaphase chromosomes when cells are cultured under conditions that inhibit DNA replication, such as low folate or in the presence of specific drugs (Sutherland and Hecht, 1985). The study of fragile sites was introduced in the 1970s by Sutherland and gained notoriety with the description of the fragile site at Xq28 (FRAXA) associated with the fragile X syndrome. Fragile sites are now generally grouped into two classes, the "rare" and the "common" fragile sites, based on their frequency of occurrence and means of induction. The rare fragile sites segregate in a Mendelian codominant manner and include the FRAXA and some 30 other fragile sites, many of which have been observed in only one or a few families. Most of the rare fragile sites are induced by conditions of folate stress (so-called folate-sensitive fragile sites), whereas others are induced by BrdU, 5-azacytidine, or the AT-specific ligands, such as distamycin A.

The common fragile sites, on the other hand, may occur on all chromosomes as a constant feature. Like the rare folate-sensitive fragile sites, some common fragile sites are also induced by folate stress and were first observed when cells from both normal individuals and those affected with the fragile X syndrome were grown under conditions to induce the FRAXA. However, unlike the rare fragile sites, they are also induced by aphidicolin and similar inhibitors of DNA replication. At present, the Genome Data Base lists 84 common fragile sites, although this number is subject to question based on criteria for inclusion. FRA3B has been of particular interest since it is the most frequent common fragile site, and is located in a chromosome band containing a putative tumor suppressor gene(s) (described below).

The cloning of the FRAXA and the discovery of trinucleotide repeat expansion in the associated *FMR1* gene marked the discovery of a new mecha-

The authors are members of The FRA3B Consortium.

*Correspondence to: Michelle M. Le Beau, Ph.D., Section of Hematology/Oncology, The University of Chicago, 5841 S. Maryland Ave, MC2115, Chicago, IL 60637.
E-mail: mmlebeau@meis.bsd.uchicago.edu

Received 28 July 1997; Accepted 1 October 1997

nism for mutagenesis in human genetic disorders. Six rare fragile sites (FRAXA, FRAXE, FRAXF, FRA16A, FRA16B, and FRA11B) have now been cloned and characterized (Warren, 1996; Yu et al., 1997). With the exception of FRA16B, all are folate-sensitive fragile sites, and the mutation leading to fragility is the expansion and methylation of a CCG repeat (expanded CAG, CTG, or GAA repeats apparently do not give rise to fragile sites) (Jalal et al., 1993). The distamycin A-inducible FRA16B was recently cloned and found to involve the expansion of a 33 bp AT-rich minisatellite repeat (Yu et al., 1997). These changes can give rise to genetic disease by modifying the expression of genes in which they reside, as is the case with FRAXA and FRAXF, or by mediating chromosome deletions, as is seen in some cases of Jacobsen syndrome. In all cases, the changes give rise to fragile sites by mechanisms that may involve formation of complex intramolecular structures that act as a block to DNA synthesis.

In addition to forming gaps or breaks on metaphase chromosomes, common fragile sites have been shown to display a number of characteristics of unstable, highly recombinogenic DNA in vitro. In particular, they are preferred sites for sister chromatid exchanges (Glover and Stein, 1987), chromosomal deletions and rearrangements (Glover and Stein, 1988; Wang et al., 1993), and the integration of transfected plasmid DNA (Rassool et al., 1991). Fragile sites have also been implicated in the initiation and perpetuation of breakage-fusion-bridge cycles, leading to gene amplification in Chinese hamster cells in vitro (Kuo et al., 1994; Coquelle et al., 1997). Based on this behavior and their location with respect to chromosomal breakpoints seen in tumor cells, it has been hypothesized that fragile sites, particularly the common fragile sites, may play a mechanistic role in the recurring chromosomal rearrangements and mutations observed in tumor cells (Le Beau and Rowley, 1984; Yunis and Soreng, 1984). These studies were based largely on the statistical correlation of fragile sites with the breakpoints of recurring abnormalities, and they made the assumption that instability at fragile sites occurs in vivo as it does in vitro. Until recently, there have been no molecular reagents to test this hypothesis directly. One common fragile site, FRA3B, has now been cloned. Perhaps surprisingly, it contains a gene (the *FHIT* gene) (Ohta et al., 1996). Perhaps not surprisingly, this gene is unstable in many tumors (described below).

The Common Fragile Site at 3p14.2, FRA3B

Sequences at FRA3B were identified using four different approaches. Wilke et al. (1996) used a positional cloning approach and fluorescence in situ hybridization (FISH) to identify a yeast artificial chromosome (YAC 850A6) that crossed FRA3B, and localized the fragile site to a large region of >100 kb that is ~160 kb telomeric to the t(3;8) breakpoint seen in a family with hereditary renal cell carcinoma. Paradee et al. (1996) mapped a series of aphidicolin-induced breakpoints in chromosome 3 containing hybrids to the same YAC and identified two clusters of breaks flanking the region characterized by Wilke et al. By cloning sequences surrounding plasmid DNA integrated after being transfected into cells treated with aphidicolin, Rassool et al. (1996) identified fragile site sequences ~350 kb distal to the t(3;8) breakpoint. Finally, Ohta et al. (1996) cloned the *FHIT* gene from the same region, and it is now known that *FHIT* spans FRA3B.

The results of these studies suggest that the common fragile sites and FRA3B in particular differ from FRAXA and other rare fragile sites in several ways. First, genomic breakage and instability occur over a very large genomic region, extending approximately from the t(3;8) breakpoint to a region at least 500 kb telomeric (Paradee et al., 1996; Rassool et al., 1996; Wilke et al., 1996). Whether this represents a single unstable region or multiple "hot spots" for DNA recombination or breakage is unknown. Second, no repeat motifs, such as CCG repeats, characteristic of most rare fragile sites, have been identified (Paradee et al., 1996; Rassool et al., 1996; Wilke et al., 1996; Boldog et al., 1997). Thus, the molecular basis for fragility at FRA3B is not yet apparent. DNA sequence analysis of 110 kb of the centromeric part of FRA3B within intron 5 of the *FHIT* gene revealed that the region is high in A–T content, and in *LINE* and *MER* repeats; however, Alu elements are reduced in frequency (Boldog et al., 1997). Several lines of evidence have suggested that fragile sites are regions of unreplicated DNA. For example, chemicals known to induce fragile sites interfere with replication fork progression and/or DNA repair (Glover et al., 1984). Le Beau et al. (unpublished data) have demonstrated recently that FRA3B is a late replicating region and that replication is further delayed, or inhibited, by aphidicolin. While late replication alone does not result in a fragile site, fragile site sequences may not be able to recover from a further delay in DNA

synthesis. Unreplicated regions of DNA would be expected to affect localized chromatin structure, and to manifest the recombinogenic properties of common fragile sites.

FHIT: A Candidate Tumor Suppressor Gene at 3p14.2

The two approaches most commonly used to determine the location of a tumor suppressor gene are karyotyping and analysis of loss of heterozygosity (LOH) of mapped polymorphic markers. Using these techniques, loss of DNA sequences on the short arm of chromosome 3 was identified in a number of different tumor types, including lung cancer (Brauch et al., 1987; Kok et al., 1987; Naylor, 1987), renal cell carcinoma (Zbar et al., 1987; Kovacs et al., 1988), breast cancer (Devilee et al., 1989; Pandis et al., 1993), and squamous cell carcinoma of the head and neck (Maestro et al., 1993; Lo et al., 1994; Wu et al., 1994; Partridge et al., 1996). At present, at least four different regions of 3p are believed to contain tumor suppressor genes: 3p12, 3p14.2, 3p21.3, and 3p25. Candidate genes have been identified at 3p14.2 (*FHIT*) and 3p25, the von Hippel Lindau (*VHL*) disease gene (Crossey et al., 1994).

The recently cloned *FHIT* gene spanning the FRA3B at 3p14.2 has generated considerable interest as a promising candidate tumor suppressor gene. *FHIT* (for fragile histidine triad) was identified by virtue of its location within a region of the genome that was homozygously deleted in epithelial cancer cell lines (Kastury et al., 1996; Ohta et al., 1996). The *FHIT* locus is composed of 10 exons (exons 5–9 are coding) distributed over ~500 kb; the gene spans the breakpoint of the constitutional t(3;8) with three 5' noncoding exons centromeric of the familial renal clear cell carcinoma-associated breakpoint, and the remaining exons telomeric. *FHIT* encodes a 1.1-kb mRNA that is ubiquitously expressed, at relatively low levels (Ohta et al., 1996). The protein has 147 amino acids (16.8 kDa) with 52% identity and 69% similarity in a core region of 109 amino acids to diadenosine 5',5'''-P^1,P^4-tetraphosphate (Ap4A) hydrolase from the fission yeast *Schizosaccharomyces pombe* (Ohta et al., 1996). Both of these proteins have sequence homology to the HIT proteins, a family of unknown function that is characterized by four conserved histidines, three of which comprise a histidine triad (the HIT sequence, HxHxH, encoded by exon 8 in the human protein).

Until recently, the function of the FHIT protein was unknown. A clue as to its possible function was provided by *S. pombe* Ap4A hydrolase, an enzyme that asymmetrically hydolyzes dinucleoside polyphosphates. In this case, Ap4A is the preferred substrate, and hydrolysis produces AMP and ATP. Barnes et al. (1996) have demonstrated recently that the FHIT protein is a typical 5',5'''-P^1,P^3-triphosphate (Ap3A) hydrolase (EC 3.6.1.29), and may function as a homodimer. Among the dinucleoside polyphosphates (ApnA, where n = 3–6), Ap3A is the preferred substrate, and AMP is always one of the reaction products. Moreover, site-directed mutagenesis demonstrated that all four conserved histidines are required for full activity, and the central histidine of the triad is essential for Ap3A hydrolase activity (Barnes et al., 1996). With the exception of FHIT and the *S. pombe* enzyme, none of the other HIT proteins are known to have Ap3A or Ap4A hydrolase activity.

What is the role of FHIT as a growth suppressor in mammalian cells? If indeed FHIT is a tumor suppressor protein, it provides the first link between dinucleotide oligophosphate metabolism and tumorigenesis. In keeping with the paradigm of tumor suppressor genes, loss of Ap3A hydrolase activity and resulting elevated levels of Ap3A and related compounds would be predicted to contribute to tumorigenesis. At present, the mechanism(s) by which this might occur is unknown. Ap3A and Ap4A have been proposed to have a variety of cellular functions, including the regulation of DNA replication and intracellular signaling in response to stress (Plateau and Blanquet, 1994); however, these hypotheses are not yet supported by strong experimental evidence. Dinucleoside oligophosphates have a variety of effects as extracellular mediators, but FHIT is apparently not an extracellular enzyme. Barnes et al. (1996) have proposed that elevated Ap3A levels stimulate a signal transduction pathway that enhances growth, or blocks a growth-inhibitory or apoptotic pathway. An example might be the inhibition of a protein kinase by elevated levels of Ap3A functioning as an ATP analog (Barnes et al., 1996).

Evidence that *FHIT* is a Tumor Suppressor Gene

The *FHIT* gene is a promising candidate tumor suppressor gene, yet several lines of evidence garnered over the past year have called into question the role of *FHIT* as a classical tumor suppressor gene. In this and the following section, we summarize the evidence in support of as well as against

FHIT as a tumor suppressor gene. The hypothesis that *FHIT* is a tumor suppressor gene is supported by three major observations: (1) *FHIT* is located in a region of the genome that is known to be rearranged and/or deleted in some tumor cells; (2) homozygous deletions have been identified in some types of tumors (gastric, lung) and, in some cases, the deletions included *FHIT* exons; and (3) aberrant *FHIT* transcripts have been identified in some tumors, particularly tumors of the digestive and respiratory tracts.

By using representational difference analysis of genomic DNA, Lisitsyn et al. (1995) identified a probe on 3p (BE758-6) that detected homozygous deletions in colon carcinoma cell lines. The probe was subsequently mapped to 3p14.2 within FRA3B. Kastury et al. (1996) defined homozygous deletions encompassing these sequences in a number of gastric, colorectal, and nasopharyngeal carcinoma cell lines, as well as LOH of an adjacent microsatellite marker (D3S1300) in uncultured gastric and colon tumors. These results raised the possibility that the same tumor suppressor gene was involved in tumors of various histological origins. *FHIT* was identified subsequently by an exon-trapping strategy using clones from a cosmid contig spanning this deleted region (Ohta et al., 1996).

More direct evidence that *FHIT* is a tumor suppressor gene was obtained by the analysis of *FHIT* transcripts in tumor cells with genomic alterations within FRA3B. Ohta et al. (1996) used reverse transcription-PCR (RT-PCR) with amplification for 30 cycles with *FHIT* primers, and reamplification with nested primers (30 cycles), to examine RNA from tumor cell lines as well as uncultured esophageal, stomach, and colon carcinomas. Aberrant transcripts were identified in 50% of the tumors studied and in most of the cell lines; the majority of the tumors and cell lines also had a normal-sized transcript (Ohta et al., 1996). Aberrant transcripts were smaller, and sequence analysis revealed the absence of various regions between exons 4–9 (exon 5 encodes the initial methionine of the open reading frame, and exon 8 encodes the HIT domain). In most of the aberrant transcripts, the beginning and end of the missing segments of the transcripts coincided with splice sites. These observations were compatible with the possibility that cDNA deletions resulted from loss of genomic regions containing the relevant lost exons. Alternatively, as described in the next section, they could also be explained by the presence of rare aberrantly spliced messages in a tissue or cell population.

If *FHIT* is the relevant tumor suppressor gene in 3p14.2, one would expect a high frequency of aberrant *FHIT* transcripts in tumors characterized by a high incidence of allele loss at 3p14, e.g., lung, breast, colon, and head and neck squamous cell carcinomas. In this regard, aberrant transcripts were detected in 30% of breast carcinomas and cell lines (Negrini et al., 1996), in 80% of primary small cell lung cancers (SCLC), and in 40% of primary nonsmall cell lung cancers (NSCLC); most of these tumors also had allelic losses for microsatellite markers within or flanking *FHIT* (Sozzi et al., 1996). In another example, Virgilio et al. (1996) detected aberrant *FHIT* transcripts in a high proportion of head and neck squamous cell carcinoma cell lines (55%), and several of these cell lines showed homozygous deletions within *FHIT* by Southern blot analysis.

In these early studies, homozygous deletions of *FHIT* exons were observed only in tumor-derived cell lines, raising concerns about the role of *FHIT* in carcinogenesis (Virgilio et al., 1996; Druck et al., 1997). Michael et al. (1997) have addressed the question of the role of *FHIT* alterations and their relationship to FRA3B in the early stages of tumor development by examining *FHIT* in a premalignant condition of the esophagus, Barrett's metaplasia, and in the associated esophageal adenocarcinomas. Using single-stage PCR, they detected aberrant *FHIT* transcripts in 12 out of 14 (86%) Barrett's metaplasias, and in 14 out of 15 (93%) of the adenocarcinomas examined from the same patients. Characterization of the altered transcripts revealed *FHIT* mRNA lacking one or more exons, with deletions of exons 5–7 being most frequent. Analysis of genomic DNA showed homozygous deletions involving exon 5 in 4 out of 20 (20%) esophageal adenocarcinomas, and hemizygous deletions in 7 cases (35%); exon 5 was shown by FISH to lie within FRA3B. Hemizygous deletions of exon 5 were detected in a few Barrett's metaplasias (no homozygous deletions were detected), indicating that genomic rearrangements within the *FHIT*/FRA3B locus may be early events, occurring in the premalignant tissue. These results demonstrate that, in some primary tumors with aberrant *FHIT* transcripts, genomic deletions of *FHIT* exons have occurred. While these data do not exclude the tumor suppressor gene model, the authors emphasize that genomic instability is a frequent characteristic of tumor cells and common fragile sites; any gene associated with these regions may thus be susceptible to such instability.

Evidence Against *FHIT* as a Tumor Suppressor Gene

Are *FHIT* deletions a cause or an effect of cancer? Since the identification of the *FHIT* gene, certain observations have cast doubt on its role as a tumor suppressor. First, it is becoming increasingly clear that the incidence of *FHIT* alterations has been overestimated in some studies due to the fact that the frequently reported "aberrant" transcripts most often represent low-abundance splicing variants observed with the use of nested PCR. Second, even in those cases in which altered transcripts clearly result from genomic alterations, these are most frequently found in only a single allele, with production of normal transcripts from the remaining allele. With some exceptions, homozygous exonic deletions appear to be rare except in cell lines. Third, mutations other than deletions are vanishingly rare. Finally, the presence of intronic deletions associated with normal transcripts in some tumors suggests an alternative hypothesis that both these deletions and those affecting exons may arise through genomic instability without providing a selective advantage to the affected cell (Thiagalingam et al., 1996; Boldog et al., 1997; Michael et al., 1997; Ong et al., 1997).

In the initial report which suggested that *FHIT* was altered in approximately 50% of gastrointestinal cancers (Ohta et al., 1996), the predominant transcript generated by RT-PCR was wild-type. Not only did the "altered" products usually represent a minority population after two rounds of PCR, but they often appeared to be of similar size. The same types of alterations were reported in a high proportion of lung cancers by Sozzi et al. (1996), who suggested that altered function of *FHIT* is critical in lung carcinogenesis. In both of these diseases, subsequent studies appear to contradict the original reports.

Thiagalingam et al. (1996), using only one round of PCR for the *FHIT* coding sequence, observed normal-sized transcripts in 29 out of 32 gastrointestinal cancers, and in each case the sequence composition was normal. Similarly, only 2 out of 16 (13%) informative cases showed loss of heterozygosity at D3S1300 (near *FHIT* exon 5) in primary gastrointestinal tumors, and none showed aberrant transcripts (Tamura et al., 1997). Of note was that 4 out of 7 (57%) cell lines had homozygous deletions of variable size, all of which included D3S1300. In another study of lung cancer cell lines, no SCLC lines (0/17) lacked *FHIT* expression or expressed only aberrant transcripts, whereas 30% (7/24) of NSCLC cell lines had these abnormalities (Yanagisawa et al., 1996). Using nonnested PCR, Fong et al. (1997) demonstrated wild-type transcripts in 98 out of the 102 lung carcinomas examined. In addition, most tumors as well as normal tissues contained a transcript that lacked exon 8, previously reported from various cDNA libraries and tumors by Boldog et al. (1997).

Van den Berg et al. (1997) observed wild-type *FHIT* transcripts in all clear-cell renal carcinoma cell lines examined; however, most of these also had aberrant transcripts. Since LOH of markers in the *FHIT* region was observed in most of the cell lines, the various transcripts observed in each sample likely originated from a single allele. With respect to primary tumors, Bugert et al. (1997) observed a breakpoint within the *FHIT* region in only 3 out of 100 sporadic nonpapillary renal cell carcinomas, whereas 94 tumors had a breakpoint proximal to *FHIT*. Normal-sized *FHIT* transcripts were detected in 33 of 34 tumors examined.

Finally, apparently identical aberrant *FHIT* transcripts have been observed in nonmalignant tissues (Boldog et al., 1997; Panagopoulos et al., 1997; van den Berg et al., 1997). Since most of the "altered" transcripts are of low abundance and are missing discrete exons in the absence of genomic exon deletions or genomic DNA splice site mutations, they should be regarded as alternative splicing variants of unknown significance. Also likely to represent splicing variants are those products containing exact insertions of novel DNA stretches between exons or exon groups.

One of the more interesting aspects of *FHIT* as a potential tumor suppressor gene concerns its chromosomal position spanning the location of the hereditary renal carcinoma associated t(3;8) (Ohta et al., 1996). *FHIT* has also been implicated in a somatic rearrangement, t(3;12), in a benign pleomorphic adenoma of the parotid gland (Geurts et al., 1997). This translocation involved the high mobility group protein gene, *HMGIC*, implicated in a variety of pleomorphic salivary gland adenomas and other benign tumors. Several other chromosomal partners have been involved in *HMGIC* rearrangements. To date, the other partner genes or sequences include a *LIM*-domain protein (LPP), the mitochondrial aldehyde dehydrogenase gene (*ALDH2*), a predicted gene with a serine/threonine-rich domain, and RTVL-H 3' LTR sequences (Kazmierczak et al., 1996; Geurts et al., 1997). In the t(3;12), a fusion *FHIT/HMGIC* transcript as well as normal *FHIT* transcripts are produced. At pres-

ent, there is no evidence demonstrating that the chimeric protein acts as a dominant negative inhibitor of the wild-type protein. Since *HMGIC* is involved in multiple translocations, the authors hypothesized that the critical feature of this rearrangement is alteration of the *HMGIC* gene, rather than *FHIT* (Geurts et al., 1997).

Tumor suppressor genes may be inactivated in at least two ways: (1) mRNA expression can be eliminated or reduced substantially by mutations in regulatory regions (as is the case with *RB1*), by gene deletions (as in *CDKN2*), or by aberrant methylation (as in *VHL*), or (2) synthesis of a functional protein may be precluded by missense, nonsense, frameshift, or splice-site gene mutations. A striking feature of *FHIT* is the apparent near absence of mutations of the second type. Several studies have examined the normal-sized transcripts by DNA sequence analysis (Negrini et al., 1996; Thiagalingam et al., 1996; Fong et al., 1997). With rare exceptions, e.g., lung carcinoma cell line H128 (Sozzi et al., 1996), there have been no demonstrable base substitutions, frame shifts, or microdeletions, as might have been expected by analogy to other tumor suppressor genes. Even homozygous deletions appear to be uncommon in most primary tumors, being seen for example in only 6 of 135 lung cancers, with only 3 of these involving a *FHIT* exon (Fong et al., 1997).

As described earlier, the 3p14.2 common fragile site, FRA3B, has been shown to represent a region of 500 kb, rather than a single breakage point (Paradee et al., 1996; Rassool et al., 1996; Wilke et al., 1996; Boldog et al., 1997). By using probes from a corresponding 300 kb contig for Southern blot analysis, Boldog et al. (1997) found homozygous deletions in nearly 90% of cervical and 50% of colorectal carcinoma cell lines. However, several deletions did not involve *FHIT* exons (also noted by Sozzi et al., 1996, and Ong et al., 1997). DNA sequence analysis of one intronic deletion failed to show additional genes, and no alterations in *FHIT* transcripts were found. Interestingly, in these samples the deletions appeared to be correlated with the presence of *TP53* mutations, suggesting they might be a consequence of genomic instability. Another observation is the presence of discontinuous deletions (Boldog et al., 1997; Druck et al., 1997; Ong et al., 1997). While the mechanism and timing of these deletions are presently unknown, their discontinuity may reflect either continued instability or the lack of clonal selection, or both. Such redundant deletions are unnecessary for the inactivation of *FHIT* and may simply reflect genetic instability (Boldog et al., 1997).

In summary, a variety of observations raise considerable doubt for the proposed role of *FHIT* as a tumor suppressor gene. We also note that, while several human tumor types have very frequent allele loss for loci within 3p14.2, the frequency of aberrant *FHIT* transcripts is considerably lower in these same tumors, which suggests that these phenomena are unrelated.

Experimental Evidence Needed to Support the Role for *FHIT* as a "Classical" Tumor Suppressor Gene

The basic criteria that have defined tumor suppressor genes are (1) the presence of loss of function mutations, (2) inactivation in both familial and sporadic tumors, and (3) the fact that the tumor phenotype can be rescued by the wild allele. Here we outline experimental evidence that is not yet available, but would provide support for a role for *FHIT* as a "classical" tumor suppressor gene. Homozygous deletions of DNA, resulting in the absence of normal *FHIT* mRNAs, were identified in a small minority of the tumors and tumor cell lines examined; these examples conform to a classical tumor suppressor model. However, the vast majority (>90%) of both the cell lines and tumor samples studied by RT-PCR show that aberrant *FHIT* transcripts are accompanied by apparently normal transcripts. Because some RT-PCR reactions can result in overrepresentation of shorter, alternatively spliced messages, it will be necessary to demonstrate that these aberrant *FHIT* transcripts result from alterations in genomic DNA sequences.

In cases where *FHIT* is expressed, albeit aberrantly, Northern blot analysis of *FHIT* mRNA should be performed to determine if mRNA expression is substantially reduced. However, for those examples where no dramatic reduction in levels of mRNA are demonstrated, it would be necessary to demonstrate that the protein product is absent or qualitatively altered. At present, there is no evidence that altered FHIT proteins are produced, i.e., no missense mutations have been identified, and the majority of aberrant transcripts are unlikely to be translated since exon 5, which contains the AUG initiator codon for protein synthesis, is lost. If altered FHIT proteins are identified and are functional, then *FHIT* may conform to a dominant negative model for tumorigenesis.

Further proof that *FHIT* is a tumor suppressor gene will necessitate demonstrating that the abolition of *FHIT* gene expression contributes to a

tumorigenic phenotype and, conversely, that reintroduction of a wild-type gene suppresses some aspect of the malignant phenotype in vitro and in vivo. Finally, the generation and phenotypic analysis of an *FHIT* knockout mouse may provide important information on *FHIT* function. In particular, an increased and earlier development of cancer in heterozygotes and in homozygotes, if viable, would provide valuable corroborative evidence that *FHIT* is a tumor suppressor gene.

Recently, we have come to appreciate that the traditional definition of tumor suppressor genes may need to be broadened, in that a number of genes have been identified that do not fulfill all of the criteria stated above. An example is the mutator genes involved in DNA mismatch repair, whose loss of function leads to genetic damage that cannot be remedied by the reintroduction of a wild-type allele. Haber and Harlow (1997) have suggested that tumor suppressor genes should be defined as "genes that sustain loss of function mutations in the development of cancer." Whether or not *FHIT* meets this simpler definition remains to be determined.

As we have argued, while *FHIT* can be considered a candidate tumor suppressor gene, additional evidence will be necessary to confirm this proposed function. There are a number of cases in which confirmation of the role of a candidate tumor suppressor gene has proved difficult. Most cases of uncertainty have resulted from difficulty in determining which of several genes within a minimal region of loss of heterozygosity or homozygous deletion may function as a tumor suppressor gene. For example, *MCC* was initially reported to be mutated in sporadic colon cancer and to be a candidate for the critical gene involved in LOH of 5q in colorectal cancer (Kinzler et al., 1991); however, subsequent results indicated that, despite properties appropriate for a tumor suppressor gene (Matsumine et al., 1996), it may rarely play a role in cancer. Selection for functional loss of the neighboring *APC* gene is instead principally responsible for LOH in the region.

A still controversial case is *DCC*, which lies within a common region of LOH on chromosome 18 and was thought for many years to be a tumor suppressor gene (Fearon et al., 1990). This has been brought into question by recent evidence that identifies *DCC* as an evolutionarily conserved receptor that directs axonal growth and by the failure of its genetic disruption to predispose to tumor formation (Fazeli et al., 1997; Leonardo et al., 1997). The closely neighboring genes, *MADR2* (*SMAD2*) and *DPC4* (*SMAD4*), have now more clearly been identified as tumor suppressor genes (Eppert et al., 1996; Hahn et al., 1996). Loss of these and possibly yet unidentified genes may largely be responsible for LOH on 18q.

The role of candidate tumor suppressor genes identified through homozygous deletions has generally proved less difficult to confirm. In rare cases, homozygous deletions have been quite large, making it difficult to be certain which of the genes affected are responsible. Deletions on 9p can be greater than 2 Mb in size (Stadler and Olopade, 1996). Despite the clear involvement of *CDKN2* (p16), the role of concomitant loss of an alternatively spliced form of the gene (with production of a second protein in a different reading frame) and of the neighboring, related gene *CDKN2B* remains uncertain in many tumor types.

The more frequently cells with deletions are generated, the smaller the selective pressure required for fixation of the mutation within the population. Deletions may be generated at high frequency in genetically unstable regions of the DNA. For example, each of the very large number of cysts developing in human autosomal dominant polycystic kidney disease type I apparently arises through independent loss of the normal allele, as a result of instability of the region of the affected *PDK1* gene (Qian et al., 1996). As a limiting case, highly unstable regions may be expected to lead to deletions at a sufficiently high rate so as to be found at significant frequency in clonal lines even in the complete absence of selection. Given the instability of the FRA3B region, perhaps compounded by genomic instability of the affected cells, we feel that the *FHIT* deletions found in tumor cell lines can perhaps be explained in this manner, without a need to postulate a tumor suppressor role for the gene. Further experimentation, as outlined above, will be required to distinguish between these two alternatives.

REFERENCES

Barnes LD, Garrison PN, Spirashvili Z, Guranowski A, Robinson AK, Ingram SW, Croce CM, Ohta M, Huebner K (1996) *FHIT*, a putative tumor suppressor in humans, is a dinucleoside 5',5'"-P1,P3-triphosphate hydrolase. Biochem 35:11529–11535.

Boldog F, Gemmill R, West J, Robinson M, Robinson L, Li E, Roche J, Todd S, Waggoner B, Lundstrom R, Jacobson J, Mullokandov MR, Klinger H, Drabkin HA (1997) Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B. Hum Mol Genet 6:193–203.

Brauch H, Johnson B, Hovis J, Yano T, Gazdar A, Pettengill OS, Graziano S, Sorenson GD, Poiesz BJ, Minna J, Linehan WM, Zbar B (1987) Molecular analysis of the short arm of chromosome 3 in

small-cell and non-small-cell carcinoma of the lung. N Engl J Med 317:1109–1113.

Bugert P, Wilhelm M, Kovacs G (1997) *FHIT* gene and the FRA3B region are not involved in the genetics of renal cell carcinomas. Genes Chromosomes Cancer 20:9–15.

Coquelle A, Pipiras E, Toledo F, Buttin G, Debatisse M (1997) Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. Cell 89:215–225.

Crossey PA, Richards FM, Foster K, Green JS, Prowse A, Latif F, Lerman MI, Zbar B, Affara N, Ferguson-Smith MA, Maher ER (1994) Identification of intragenic mutations in the von Hippel-Lindau disease tumour suppressor gene and correlation with disease phenotype. Hum Mol Genet 3:1303–1308.

Devilee P, van den Broek M, Kuipers-Dijkshoorn N, Kolluri R, Khan PM, Pearson PL, Cornelisse CJ (1989) At least four different chromosomal regions are involved in loss of heterozygosity in human breast carcinoma. Genomics 5:554–560.

Druck T, Hadaczek P, Fu TB, Ohta M, Spirashvili Z, Baffa R, Negrini M, Kastury K, Veronese ML, Rosen D, Rothstein J, McCue P, Cotticelli MG, Inoue H, Croce CM, Huebner K (1997) Structure and expression of the human *FHIT* gene in normal and tumor cells. Cancer Res 57:504–512.

Eppert K, Scherer SW, Ozcelik H, Pirone R, Hoodless P, Kim H, Tsui LC, Bapat B, Gallinger S, Andrulis IL, Thomsen GH, Wrana JL, Attisano L (1996) *MADR2* maps to 18q21 and encodes a TGFβ-regulated MAD-related protein that is functionally mutated in colorectal carcinoma. Cell 86:543–552.

Fazeli A, Dickinson SL, Hermiston ML, Tighe RV, Steen RG, Small CG, Stoeckli ET, Keino-Masu K, Masu M, Rayburn H, Simons J, Bronson RT, Gordon JI, Tessier-Lavigne M, Weinberg RA (1997) Phenotype of mice lacking functional Deleted in colorectal cancer (*DCC*) gene. Nature 386:796–804.

Fearon ER, Cho KR, Nigro JM, Kern SE, Simons JW, Ruppert JM, Hamilton SR, Preisinger AC, Thomas G, Kinzler KW, Vogelstein B (1990) Identification of a chromosome 18q gene that is altered in colorectal cancers. Science 247:49–56.

Fong KM, Biesterveld EJ, Virmani A, Wistuba I, Sekido Y, Bader SA, Ahmadian M, Ong ST, Rassool FV, Zimmerman PV, Giaccone G, Gazdar AF, Minna JD (1997) *FHIT* and FRA3B 3p14.2 allele loss are common in lung cancer and preneoplastic bronchial lesions and are associated with cancer-related *FHIT* cDNA splicing aberrations. Cancer Res 57:2256–2267.

Geurts JMW, Schoenmakers EFPM, Röijer E, Stenman G, Van de Ven WJM (1997) Expression of reciprocal hybrid transcripts of *HMGIC* and *FHIT* in a pleomorphic adenoma of the parotid gland. Cancer Res 57:13–17.

Glover TW, Stein CK (1987) Induction of sister chromatid exchanges at common fragile sites. Am J Hum Genet 41:882–890.

Glover TW, Stein CK (1988) Chromosome breakage and recombination at fragile sites. Am J Hum Genet 43:265–273.

Glover TW, Berger C, Coyle J, Echo B (1984) DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. Hum Genet 67:136–142.

Haber D, Harlow E (1997) Tumour-suppressor genes: Evolving definitions in the genomic age. Nat Genetics 16:320–322.

Hahn SA, Schutte M, Hoque AT, Moskaluk CA, da Costa LT, Rozenblum E, Weinstein CL, Fischer A, Yeo CJ, Hruban RH, Kern SE (1996) *DPC4*, a candidate tumor suppressor gene at human chromosome 18q21.1. Science 271:350–353.

Jalal SM, Lindor NM, Michels VV, Buckley DD, Hoppe DA, Sarkar G, Dewald GW (1993) Absence of chromosome fragility at 19q13.3 in patients with myotonic dystrophy. Am J Med Genet 46:441–443.

Kastury K, Baffa R, Druck T, Ohta M, Cotticelli MG, Inoue H, Negrini M, Rugge M, Huang D, Croce CM, Palazzo J, Huebner K (1996) Potential gastrointestinal tumor suppressor locus at the 3p14.2 FRA3B site identified by homozygous deletions in tumor cell lines. Cancer Res 56:978–983.

Kazmierczak B, Pohnke Y, Bullerdiek J (1996) Fusion transcripts between the *HMGIC* gene and RTVL-H-related sequences in mesenchymal tumors without cytogenetic aberrations. Genomics 38:223–226.

Kinzler KW, Nilbert MC, Vogelstein B, Bryan TM, Levy DB, Smith KJ, Preisinger AC, Hamilton SR, Hedge P, Markham A, Carlson M, Joslyn G, Groden J, White R, Miki Y, Miyoshi Y, Nishisho I, Nakamura Y (1991) Identification of a gene located at chromosome 5q21 that is mutated in colorectal cancers. Science 251:1366–1370.

Kok K, Osinga J, Carritt B, Davis MB, van der Hout AH, van der Veen AY, Landsvater RM, de Leij LF, Berendsen HH, Postmus PE, Poppema S, Buys CHCM (1987) Deletion of a DNA sequence at the chromosomal region 3p21 in all major types of lung cancer. Nature 330:578–581.

Kovacs G, Erlandsson R, Boldog F, Ingvarsson S, Muller-Brechlin R, Klein G, Sumegi J (1988) Consistent chromosome 3p deletion and loss of heterozygosity in renal cell carcinoma. Proc Natl Acad Sci USA 85:1571–1575.

Kuo MT, Vyas RC, Jiang LX, Hittelman WN (1994) Chromosome breakage at a major fragile site associated with P-glycoprotein gene amplification in multidrug-resistant CHO cells. Mol Cell Biol 14:5202–5211.

Le Beau MM, Rowley JD (1984) Heritable fragile sites in cancer. Nature 308:607–608.

Leonardo ED, Hinck L, Masu M, Keino-Masu K, Ackerman SL, Tessier-Lavigne M (1997) Vertebrate homologues of C. elegans UNC-5 are candidate netrin receptors. Nature 386:833–838.

Lisitsyn NA, Lisitsina NM, Dalbagni G, Barker P, Sanchez CA, Gnarra J, Linehan WM, Reid BJ, Wigler MH (1995) Comparative genomic analysis of tumors: Detection of DNA losses and amplification. Proc Natl Acad Sci USA 92:151–155.

Lo K-W, Tsao S-W, Huang DP, Lee JCK (1994) Detailed deletion mapping on the short arm of chromosome 3 in nasopharyngeal carcinomas. Int J Oncol 4:1359–1364.

Maestro R, Gasparotto D, Vukosavljevic T, Barzan L, Sulfano S, Boiocchi M (1993) Three discrete regions of deletion at 3p in head and neck cancers. Cancer Res 53:5775–5779.

Matsumine A, Senda T, Baeg GH, Roy BC, Nakamura Y, Noda M, Toyoshima K, Akiyama T (1996) *MCC*, a cytoplasmic protein that blocks cell cycle progression from the G0/G1 to S phase. J Biol Chem 271:10341–10346.

Michael D, Beer DG, Wilke CM, Miller DE, Glover TW (1997) Frequent deletions of *FHIT* and FRA3B in Barrett's metaplasia and esophageal adenocarcinomas. Oncogene, 15:1653–1659.

Naylor SL, Johnson BE, Minna JD, Sakaguchi AY (1987) Loss of heterozygosity of chromosome 3p markers in small-cell lung cancer. Nature 329:451–454.

Negrini M, Monaco C, Vorechovsky I, Ohta M, Druck T, Baffa R, Huebner K, Croce CM (1996) The *FHIT* gene at 3p14.2 is abnormal in breast carcinomas. Cancer Res 56:3173–3179.

Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, Palazzo J, Siprashvili Z, Mori M, McCue P, Druck T, Croce CM, Huebner K (1996) The *FHIT* gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. Cell 84:587–597.

Ong ST, Fong KM, Bader SA, Minna JD, Le Beau MM, McKeithan TW, Rassool FV (1997) Precise localization of the *FHIT* gene to the common fragile site at 3p14.2 (FRA3B), and characterization of homozygous deletions within FRA3B that affect *FHIT* transcription in tumor cell lines. Genes Chromosomes Cancer 20:16–23.

Panagopoulos I, Thelin S, Mertens F, Mitelman F, Aman P (1997) Variable *FHIT* transcripts in non-neoplastic tissues. Genes Chromosomes Cancer 19:215–219.

Pandis N, Jin Y, Limon J, Bardi G, Idvall I, Mandahl N, Mitelman F, Heim S (1993) Interstitial deletion of the short arm of chromosome 3 as a primary chromosome abnormality in carcinomas of the breast. Genes Chromosomes Cancer 6:151–155.

Paradee W, Wilke CM, Hoge A, Glover TW, Smith DI (1996) A 350-kb cosmid contig in 3p14.2 that crosses the t(3;8) hereditary renal cell carcinoma translocation breakpoint and 17 aphidicolin-induced FRA3B breakpoints. Genomics 35:87–93.

Partridge M, Emilion G, Langdon JD (1996) LOH at 3p correlates with a poor survival in oral squamous cell carcinoma. Br J Cancer 73:366–371.

Plateau P, Blanquet S (1994) Dinucleoside oligophosphates in micro-organisms. Adv Microbiol Physiol 36:82–109.

Qian F, Watnick TJ, Onuchic LF, Germino GG (1996) The molecular basis of focal cyst formation in human autosomal dominant polycystic kidney disease type I. Cell 87:979–987.

Rassool FV, McKeithan TW, Neilly ME, van Melle E, Espinosa III R, Le Beau MM (1991) Preferential integration of marker DNA into the chromosomal fragile site at 3p14: An approach to cloning fragile sites. Proc Natl Acad Sci USA 88:6657–6661.

Rassool FV, Le Beau MM, Shen M-L, Neilly ME, Espinosa III R, Ong ST, Boldog F, Drabkin H, McCarroll R, McKeithan TW (1996) Direct cloning of DNA sequences from the common fragile site region at chromosome band 3p14.2. Genomics 35:109–117.

Sozzi G, Veronese ML, Negrini M, Baffa R, Cotticelli MG, Inoue H, Tornielli S, Pilotti S, DeGregorio L, Pastorino U, Pierotti MA,

Ohta M, Huebner K, Croce CM (1996) The *FHIT* gene at 3p14.2 is abnormal in lung cancer. Cell 85:17–26.

Stadler WM, Olopade OI (1996) The 9p21 region in bladder cancer cell lines: Large homozygous deletion inactivate the *CDKN2*, *CDKN2B* and *MTAP* genes. Urol Res 24:239–244.

Sutherland GR, Hecht F (1985) Fragile Sites on Human Chromosomes. Oxford: Oxford University Press.

Tamura G, Sakata K, Nishizuka S, Maesawa C, Suzuki Y, Iwaya T, Terashima M, Saito K, Satodate R (1997) Analysis of the fragile histidine triad gene in primary gastric carcinomas and gastric carcinoma cell lines. Genes Chromosomes Cancer 20:98–102.

Thiagalingam S, Lisitsyn NA, Hamaguchi M, Wigler MH, Willson JKV, Markowitz SD, Leach FS, Kinzler KW, Vogelstein B (1996) Evaluation of the *FHIT* gene in colorectal cancers. Cancer Res 56:2936–2939.

Van den Berg A, Draaijers TG, Kok K, Timmer T, Van der Veen AY, Veldhuis PMJF, de Leij L, Gerhartz CD, Naylor SL, Smith DI, Buys CHCM (1997) Normal *FHIT* transcripts in renal cell cancer and lung cancer-derived cell lines, including a cell line with homozygous deletion in the FRA3B region. Genes Chromosomes Cancer 19:220–227.

Virgilio L, Shuster M, Gollin SM, Veronese ML, Ohta M, Huebner K, Croce CM (1996) *FHIT* gene alterations in head and neck squamous cell carcinomas. Proc Natl Acad Sci USA 93:9770–9775.

Wang ND, Testa JR, Smith DI (1993) Determination of the specificity of aphidicolin-induced breakage of the human 3p14.2 fragile site. Genomics 17:341–347.

Warren ST (1996) The expanding world of trinucleotide repeats. Science 271:1374–1375.

Wilke CM, Hall BK, Hoge A, Paradee W, Smith DI, Glover TW (1996) FRA3B extends over a broad region and contains a spontaneous HPV16 integration site: Direct evidence for the coincidence of viral integration sites and fragile sites. Hum Mol Genet 5:187–195.

Wu CL, Sloan P, Read AP, Harris RH, Thakker NS (1994) Deletion mapping on the short arm of chromosome 3 in squamous cell carcinoma of the oral cavity. Cancer Res 54:6484–6488.

Yanagisawa K, Kondo M, Osada H, Uchida K, Takagi K, Masuda A, Takahashi T, Takahashi T (1996) Molecular analysis of the *FHIT* gene at 3p14.2 in lung cancer cell lines. Cancer Res 56:5579–5582.

Yu S, Mangelsdorf M, Hewett D, Hobson L, Baker E, Eyre HJ, LapsysN, Le Paslier D, Doggett NA, Sutherland GR, Richards RI (1997) Human chromosomal fragile site FRA16B is an amplified AT-rich minisatellite repeat. Cell 88:367–374.

Yunis JJ, Soreng AL (1984) Constitutive fragile sites and cancer. Science 226:1199–1204.

Zbar B, Brauch H, Talmadge C, Linehan WM (1987) Loss of alleles of loci on the short arm of chromosome 3 in renal cell carcinoma. Nature 327:721–724.

m/s SYNTHÈSE

médecine/sciences 1998 ; 14 : 283-90

Chromosome 3p, gènes suppresseurs de tumeurs et gènes de sémaphorines en 3p21.3

Joëlle Roche
James West
Robert Gemmill
Harry Drabkin

Des anomalies affectent le chromosome 3p humain dans différents types de cancer, dont les cancers pulmonaires. Trois régions de délétions chromosomiques sont observées en 3p25, 3p21.3 et 3p12-14 et la fréquence élevée de perte d'hétérozygotie dans les tumeurs suggère la présence d'un ou de plusieurs gènes suppresseurs de tumeurs dans ces régions. Si un gène suppresseur, le gène *VHL (von Hippel Lindau)*, a été localisé en 3p25, la présence d'autres gènes suppresseurs en 3p21.3 et 3p12-14 n'a pas encore été démontrée. La région 3p14.2 inclut le site de la translocation t(3;8) associée au cancer héréditaire du rein, le site fragile commun *FRA3B* et le gène suppresseur de tumeur supposé *FHIT*, mais les délétions observées de façon récurrente pourraient résulter d'une instabilité génomique de la région. Dans la région 3p21.3, sont localisés le gène *GNAI2* codant pour une sous-unité d'une protéine G inhibitrice et le gène de la sémaphorine IV. Outre leur rôle dans le guidage des axones, les sémaphorines pourraient être impliquées dans la différenciation et la reconnaissance cellulaires.

Le cancer du poumon est l'une des premières causes de mortalité dans les pays développés. Un certain nombre d'altérations génétiques affectent des proto-oncogènes tel *MYC* ainsi que des gènes suppresseurs de tumeurs comme *RB* et *P53*. En outre, d'autres anomalies impliquent des délétions et des amplifications chromosomiques (pour mise au point récente, *voir* [1]) : des études cytogénétiques et des analyses de perte d'hétérozygotie (LOH) ont mis en évidence des délétions affectant la région 3p dans des lignées cellulaires provenant de carcinomes du poumon à petites cellules (SCLC), ces altérations étant un événement précoce dans le développement des tumeurs (pour revue, *voir* [2]).

Les tumeurs de phénotype SCLC mettent en jeu des cellules pulmonaires d'origine épithéliale qui synthétisent des marqueurs de différenciation neuroendocrine comme l'énolase spécifique des neurones. La fréquence de perte d'hétérozygotie

ADRESSES

J. Roche : *professeur à l'Université de Poitiers.* IBMIG, Université de Poitiers, 40, avenue du Recteur-Pineau, 86022 Poitiers Cedex, France. J. West : *Ph.D., biologiste informaticien.* R. Gemmill : *associate professor.* H. Drabkin : *associate professor.* University of Colorado, health sciences center, Division of medical oncology, Box B171, 4200 East Ninth Avenue, Denver, CO 80262, États-Unis.

de la région 3p atteint pratiquement 100 % des cas; cette fréquence est réduite à environ 75 % dans le cas de tumeurs pulmonaires qui ne sont pas à petites cellules, NSCLC [3, 4]. Cette perte d'hétérozygotie est principalement localisée dans 3 régions: 3p25, 3p21.3 et 3pcen-3p14 *(figure 1)*. Ces observations ont conduit à l'hypothèse qu'un ou plusieurs gènes suppresseurs de tumeurs résident sur le chromosome 3p [2].

Dans le cas de la région 3p25, la présence du gène suppresseur *VHL (von Hippel Lindau)* [5], a été clairement montrée. Ce gène code pour un effecteur négatif de l'élongation de la transcription [6] et son inactivation prédispose les individus à différents cancers dont le carcinome rénal.

Dans le cas des régions 3p21.3 et 3pcen-3p14, le transfert d'ADN provenant de ces régions permet de réduire l'apparition de tumeurs chez des souris immunodéficientes [7, 8]. Cependant, les nombreuses données expérimentales n'ont pas encore permis de caractériser sans ambiguïté un gène suppresseur [2].

Cet article a pour objectif de présenter une mise au point des connaissances concernant la recherche de gènes suppresseurs de tumeurs sur 3p. Dans une première partie, nous décrirons les caractéristiques des régions 3pcen-3p14 et 3p21.3, l'accent sera plus particulièrement mis sur la description de la région 3p21.3. Ensuite, nous exposerons les propriétés des sémaphorines dont deux gènes sont localisés en 3p21.3. Ces gènes étant inclus dans une région qui présente un effet antitumoral, nous discuterons leur participation éventuelle au développement des tumeurs.

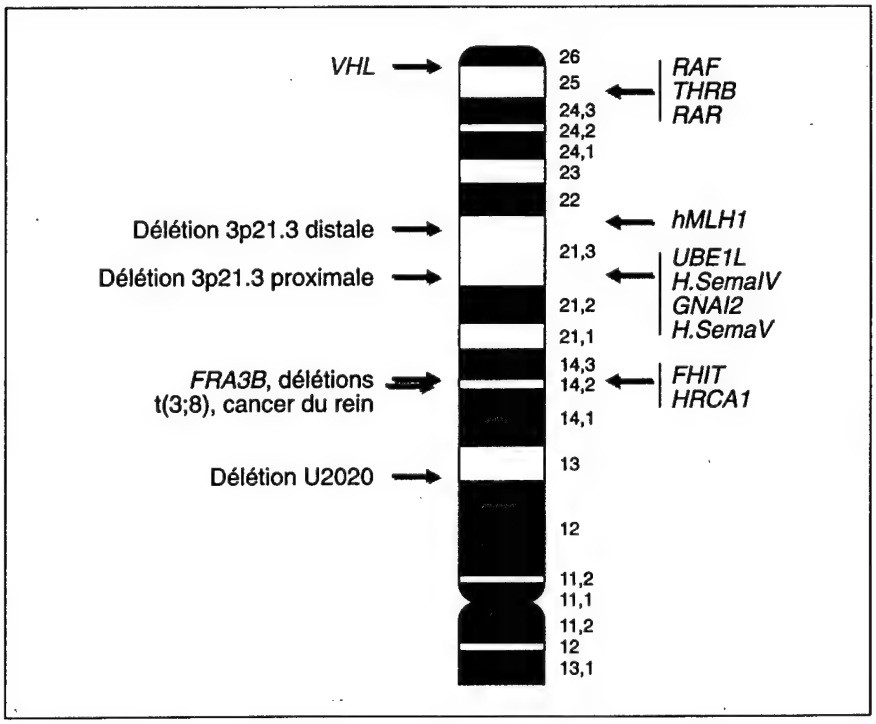


Figure 1. ***Chromosome 3p humain.*** *Les régions de délétions homozygotes sont indiquées sur la partie gauche de la figure. La première délétion homozygote a été décrite dans la lignée SCLC, U2020 [45]. Les gènes candidats impliqués dans le développement de tumeurs sont indiqués sur la partie droite [1]:* RAF *(gène de la protéine-kinase Raf),*THRB *(gène du récepteur β de l'hormone thyroïdienne),* RAR *(gène du récepteur de l'acide rétinoïque),* hMLH1 *(gène d'une protéine de réparation de l'ADN),* UBE1L *(ou D8, code pour une enzyme d'activation de l'ubiquitine),* SemaIV *et* SemaV *(gènes de sémaphorines),* GNAI2 *(gène de la sous-unité de la protéine Gi2),* FHIT *(gène de la pyrophospho-hydrolase diadénosine tétraphosphate),* HRCA1 (hereditary renal carcinoma associated-1 gene). *Le gène* VHL (von Hippel Lindau) *est le seul pour lequel la fonction de suppresseur de tumeurs soit clairement établie.*

Cartographie des régions 3p14.2 et 3p21.3

Région 3p14.2

La perte d'hétérozygotie de la région 3p12-14 est fréquente dans les cas de carcinomes du rein, du col de l'utérus, de la gorge et du nasopharynx. Cette région complexe inclut le site de la translocation t(3;8) (p14.2;q24.1) impliquée dans le cancer héréditaire du rein [9, 10], ainsi que le site fragile commun *FRA3B* [11]. En fait, ce site fragile représente une région d'environ 200 kb plutôt qu'un point unique de fragilité. En effet, il contient des sites de cassure induits par l'aphidicoline ainsi que les sites d'intégration du vecteur pSV2Néo et du virus du papillome humain HPV16 [2, 12].

Nous avons pu mettre en évidence, dans la région du marqueur *D3S1300* et à sa proximité, une série de délétions homozygotes dans des lignées cellulaires provenant de carcinomes du col de l'utérus et du côlon et, à une fréquence moindre, dans des lignées de carcinomes du poumon et du sein [12]; la taille de la plus courte séquence d'ADN délétée est d'environ 40 kb. Pendant notre étude, le gène *FHIT (fragile histidine triad)* a été cloné et localisé de part et d'autre du site de translocation t(3;8) [13]. La protéine codée par ce gène présente 50 % d'analogie avec le produit du gène *Aph1p* de levure. Ce dernier code pour une pyrophospho-hydrolase diadénosine tétraphosphate dont la fonction est de cliver de façon asymétrique un dinucléotide, la diadénosine 5', 5'''-P^1,P^4-tétraphosphate (Ap_4A), en ATP et AMP. La fonction de Ap_4A n'est pas clairement établie, mais il semble que cette molécule soit impliquée dans la réplication et la réparation de l'ADN ainsi que dans la prolifération cellulaire. Le gène *FHIT*, composé d'au moins 10 exons, s'étend sur 500 kb; seul l'exon 5 de ce gène est localisé dans la région de délétions homozygotes que nous avons décrites [12]. Ohta *et al.* ont montré

que l'expression de ce gène était altérée dans environ 50 % des cas de carcinomes de l'œsophage, de l'estomac et du côlon et ils ont émis l'hypothèse que *FHIT* serait un gène suppresseur de tumeurs [13].
Cependant, un certain nombre d'observations allant à l'encontre de cette hypothèse [2, 12], des études *in vivo* sont nécessaires pour montrer le rôle de *FHIT* dans le développement des tumeurs, au niveau de la croissance ou de la différenciation cellulaires.
Le séquençage de l'ADN a mis en évidence une analogie entre la teneur et la composition en éléments répétés d'une région de 110,4 kb de *FRA3B* et celles de la région de 152 kb du chromosome X, impliquée dans le syndrome de l'X fragile [12]. Cette dernière contient le site fragile *FRAXA*. De même, la région impliquée dans le syndrome de l'X fragile et le site fragile *FRA3B* sont riches en AT (63 %) et pauvres en gènes. L'hypothèse de la présence d'un gène suppresseur en 3p14 avait été initialement proposée. Mais la présence et la caractérisation des délétions en 3p14 suggèrent plutôt une instabilité génomique de cette région que la perte sélective d'un gène suppresseur. En effet, cette hypothèse expliquerait l'existence de délétions discontinues dans des tumeurs et la présence de délétions dans des échantillons non tumoraux.

Carte physique de la région 3p21.3

La région 3p21.3 présente des délétions homozygotes dans des lignées SCLC; on la subdivise en deux sous-régions, proximale et distale. La carte physique de la région proximale a été réalisée à partir de l'ADN de trois lignées cellulaires provenant de carcinomes du poumon à petites cellules, NCI-H740 [14], GLC20 [15] et NCI-H1450 [16]. L'ADN de ces trois lignées cellulaires présente une délétion homozygote dans la région proche du marqueur *D3F15S2*, ce qui conduit à la délétion du gène *GNAI2* codant pour la sous-unité Gαi2 d'une protéine G hétérotrimérique [17]. Les bordures télomériques de ces trois délétions sont localisées près d'un point de cassure se produisant spontanément dans l'hybride somatique de radiation 2A3-CT, et cette région est particulièrement instable. Leurs extrémités centromériques se situent près du point de la translocation t(3;7) impliquée dans le syndrome de Greig *(figure 2A)*. La taille de la région commune délétée dans les trois lignées étudiées est d'environ 450 kb, elle correspond à la délétion de la lignée GLC20 *(figure 2B)*.
Des délétions homozygotes dans cette région ont également été mises en évidence directement dans des tumeurs du poumon et non plus uniquement dans des lignées cellulaires en culture [4].
La zone de délétion distale a été observée dans cinq lignées provenant de cancers pulmonaires [18], puis dans deux tumeurs du poumon [16]. De taille estimée à 800 kb [18], elle est située en 3p21.3-p22 à 15 cM de la délétion impliquant le gène *GNAI2* [16]. Des études complémentaires sont nécessaires pour définir ses limites de façon plus précise. Il est intéressant de remarquer que le gène de réparation de l'ADN, *hMLH1* [19] impliqué dans le cancer du côlon, est situé à proximité immédiate de cette délétion distale. Cependant, nous n'avons pas pu mettre en évidence des remaniements importants de ce gène dans 24 lignées tumorales et 25 tumeurs pulmonaires testées [16].

Gènes de la région 3p21.3

Plus de 50 gènes ont été cartographiés en 3p21 dont plusieurs codant pour des protéines à doigts de zinc. Aucun gène n'a pour l'instant été décrit dans la région de délétion 3p21.3 distale, mais on en a localisé plusieurs dans la région 3p21.3 proximale [2]. Nous ne décrirons ici que les gènes présents dans la région délétée commune aux trois lignées SCLC décrites.
Le gène *GNAI2* code pour la protéine Gαi2, l'une des trois isoformes de Gαi, responsable de l'inhibition de l'activité adényl cyclasique et de la stimulation de canaux potassiques dépendant de l'ATP. Cette protéine est aussi impliquée dans des phénomènes plus complexes comme la prolifération des fibroblastes ou la stimulation de MAP-kinases. Le gène *GNAI2* pourrait donc représenter un bon candidat comme gène suppresseur de cancer. A l'encontre de cette hypothèse, Daly *et al.* [14] ont montré par RT PCR, que ce gène était exprimé dans 14 lignées cellulaires provenant de cancer du poumon mais aucune étude de mutation ponctuelle n'a été réalisée par ces auteurs dans ces lignées. Cependant, des mutations qui modifient l'activité GTPase ont été décrites dans des tumeurs de la thyroïde et de l'ovaire. En outre, des souris rendues déficientes en Gi2 développent une maladie inflammatoire du gros intestin qui évolue vite vers la colite ulcéreuse et l'adénocarcinome [20]. Le développement d'un adénocarcinome pourrait cependant résulter des processus inflammatoires et les résultats devraient être comparés à ceux obtenus dans des conditions d'élevage stérile des animaux.
Des expériences de capture d'exons *(exon trapping)* nous ont permis d'isoler un ADNc que nous avons appelé *SemaIV* et qui code pour une nouvelle protéine de la famille des sémaphorines/collapsines [16], famille dont les caractéristiques et les fonctions seront décrites dans le prochain paragraphe. Le même ADNc a été cloné et nommé *SEMA3F* par le groupe de Naylor [21]. En outre, dans cette même région, un second gène de la famille des sémaphorines *(SemaV)* a été cloné par le groupe de Minna [22]. Le gène *GNAT-1*, qui code pour la transducine, une protéine fixant le GTP exprimée uniquement dans la rétine, a été localisé à proximité [22]. La région minimale de délétion en 3p21.3 contient donc une zone de duplication où, dans un fragment de 70 kb, deux gènes codant pour des protéines fixatrices de GTP sont flanqués de deux gènes codant pour des sémaphorines.
Une carte transcriptionnelle de la région 3p21.3 a été récemment publiée à partir de la construction d'un *contig* de cosmides de 600 kb [23] où 15 séquences potentiellement transcrites et 18 îlots CpG ont été positionnés. Cette carte confirme la présence et les positions relatives des deux gènes *SemaIV* et *SemaV* ainsi que des deux gènes *GNAI2* et *GNAT-1*. Un autre ADNc *(LuCa)*, qui code pour une protéine dont la fonction reste à définir, est situé dans la région commune et minimale de délétion.
Au cours des derniers mois, le *Genome Center* de l'Université de Washington (Saint Louis, MI, USA) a engagé un programme de séquen-

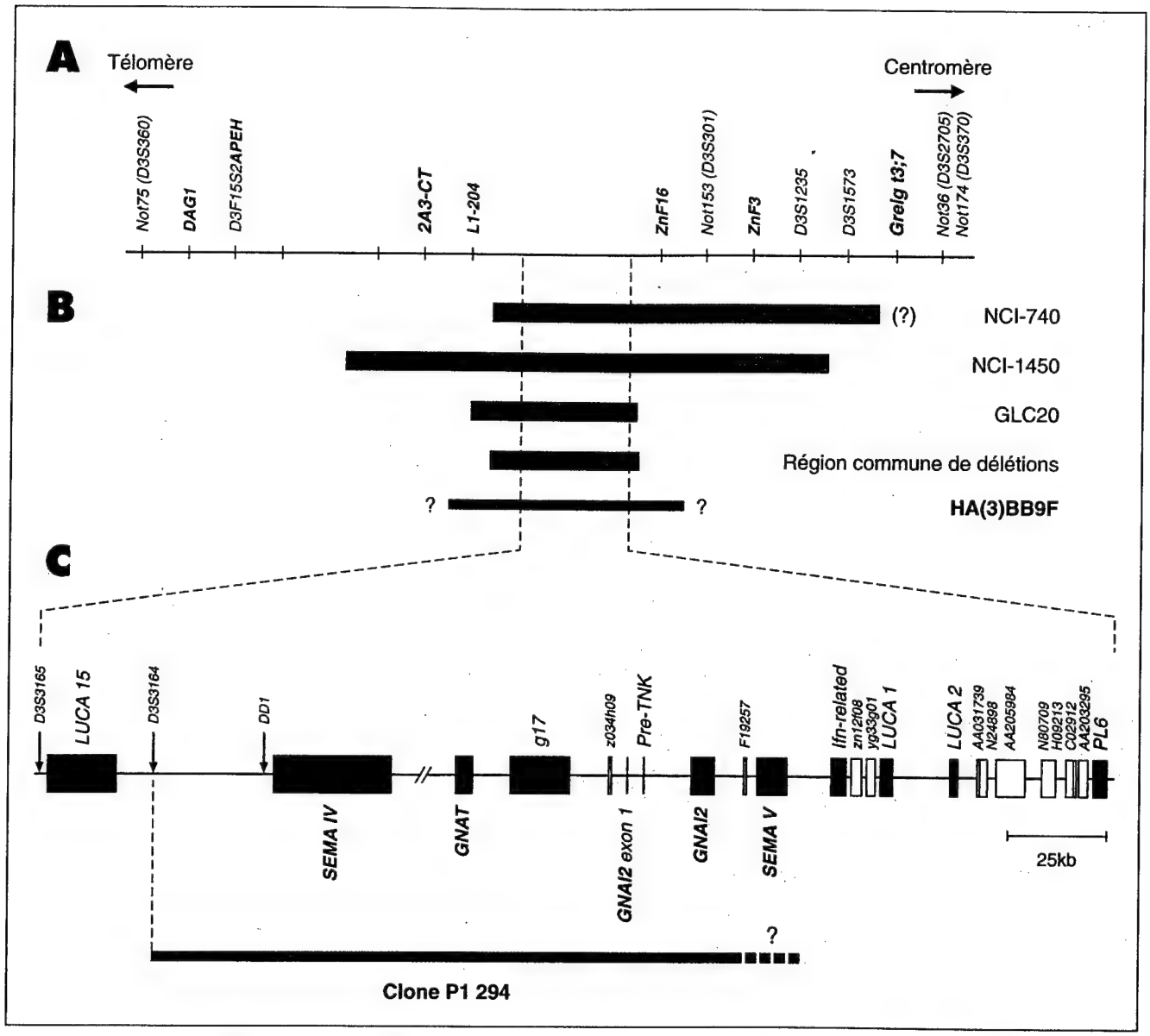


Figure 2. ***Carte de la région proximale 3p21.3. A.*** *Carte physique de la région.* ***B.*** *Représentation des délétions homozygotes des lignées SCLC (NCI-740, NCI-1450 et GLC20), de la région commune de délétion et du fragment de 2 Mb de la lignée hybride, HA(3)BB9F, possédant l'activité antitumorale. Les points d'interrogation représentent l'incertitude au niveau des limites des délétions.* ***C.*** *Carte transcriptionnelle de la région commune de délétions. Les séquences ont été obtenues à partir du serveur du département de génétique de l'Université de Washington (ftp://genome.wustl.edu/pub/gsc1/sequence/st.louis/human/) et comparées aux séquences de Genbank (nr et EST) au moyen de BLASTn: séquences transcrites identifiées (rectangles rouges) et groupes d'ADNc non identifiés (rectangles blancs). Le gène* g17 *code pour un transporteur (U49082). La localisation du clone P1 294 à l'effet antitumoral est indiquée dans la dernière partie de la figure; elle est déduite des indications fournies par les travaux du groupe de Naylor [24].*

çage systématique de la région 3p21. L'accès aux séquences données dans les banques nous a permis d'établir la carte où sont localisés un certain nombre de gènes et de séquences potentiellement transcrites *(figure 2C)*. Une publication récente du groupe de Naylor [24] apporte des arguments forts concernant la présence éventuelle d'un gène suppresseur de tumeurs dans la région 3p21.3. En effet, ces auteurs ont pu réduire la région d'intérêt de 2 Mb à un fragment d'ADN de 80 kb cloné dans un vecteur P1 *(figure 2C)* et ils ont montré que ce fragment était suffisant pour la suppression de la croissance tumorale *in vivo*. Selon la carte présentée par les auteurs, les gènes *SemaIV* et *GNAI2* seraient présents dans ce fragment de 80 kb [2, 24]. Il nous semble donc intéressant de présenter les caractéristiques et les propriétés des sémaphorines.

Les sémaphorines/ collapsines

Ces protéines sont exprimées dans de nombreuses espèces, des insectes aux mammifères. Ce sont, soit des protéines de surface membranaires, soit des protéines excrétées [25-27]. Elles présentent des caractéristiques communes : une taille d'environ 750 acides aminés, un domaine conservé appelé domaine « Sema » d'environ 500 acides aminés caractérisé par la présence de 14 à 16 cystéines, plusieurs blocs d'acides aminés conservés et un site potentiel d'amino-glycosylation. Elles sont classées en cinq familles en fonction de la présence d'un domaine de type immunoglobuline (69 acides aminés), de motifs thrombospondine et/ou d'un domaine transmembranaire *(figure 3)*. Des analogies de séquences au niveau du domaine « Sema » ont été trouvées dans des protéines virales d'un virus herpès [28] ainsi que chez des poxvirus (virus de la vaccine et de la variole) [25] ; les sémaphorines des poxvirus ne possèdent qu'une forme tronquée de ce domaine. On estime qu'il existe une vingtaine de sémaphorines différentes chez l'homme.

Les sémaphorines IV et V appartiennent à la classe III (présence d'un domaine de type immunoglobuline et séquence carboxy-terminale riche en acides aminés basiques). L'analogie la plus forte est celle qui existe entre la protéine SemaIV et la sémaphorine murine SemE, les protéines SemaIIID (humaine et murine) et leur équivalent chez le poulet, la collapsine I *(figure 4)*. Ces sémaphorines sont excrétées et interviennent au cours du développement dans la croissance de certains neurones et leur guidage : ce sont des effecteurs présentant un effet répulsif sur le guidage des cônes de croissance ([29, 30] et revues [26, 27, 31-33]).

Quelle est la fonction de la protéine SemaIV ?

Pour l'instant, nous ne la connaissons pas. Son gène est exprimé sous la forme d'un transcrit qui subit un épissage alternatif ; la forme majoritaire de 3 kb est présente dans plusieurs tissus humains fœtaux et adultes, notamment le poumon. Il semble peu probable que cette protéine soit impliquée uniquement dans des fonctions neuronales ; elle pourrait l'être également dans des phénomènes de différenciation, de reconnaissance, d'adhérence et/ou de migration cellulaires. Différentes observations vont d'ailleurs dans le sens d'un rôle plus large des sémaphorines : (1) on dénombre chez l'homme environ une vingtaine de sémaphorines ; (2) l'invalidation par recombinaison homologue du gène *SemaIII* chez la souris a montré que la sémaphorine III est nécessaire non seulement au développement du tissu nerveux mais aussi à celui des os et du cœur [34] ; (3) les motifs thrombospondine présents dans les sémaphorines de classe V entrent dans la composition de la matrice extracellulaire [35] ; (4) la présence de protéines contenant le motif Sema chez des virus suggère qu'elles auraient une fonction dans l'immunosuppression [25] ; (5) la sémaphorine CD100 induit l'agrégation des lymphocytes B et améliore leur viabilité *in vitro* [32, 36] ; cette sémaphorine jouerait donc un rôle dans le fonctionnement du système immunitaire ; (6) le rôle de la sémaphorine E humaine dans le développement de la résistance aux drogues (non MDR) dans les cancers vient d'être rapporté par Yamada *et al.* [37].

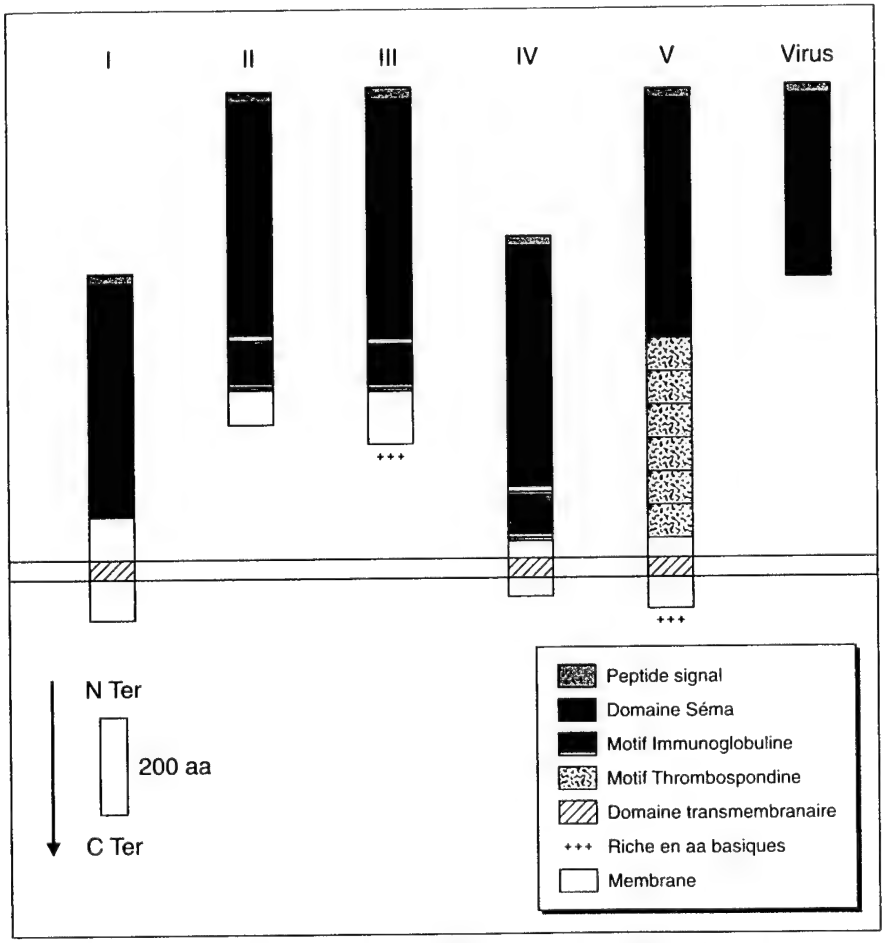


Figure 3. ***Représentation des cinq familles de sémaphorines et de la forme virale des poxvirus.*** *Le schéma publié par Püschel* et al. *[27] a été adapté en y incluant la forme virale des poxvirus. Les différents domaines sont représentés à l'échelle et les protéines sont localisées par rapport à la membrane cellulaire : les formes II, III et virales sont sécrétées et les formes I, IV et V sont membranaires. Les extrémités carboxy-terminales des formes membranaires sont cytoplasmiques. aa : acides aminés, N-Ter : extrémité amino-terminale, C-ter : extrémité carboxy-terminale.*

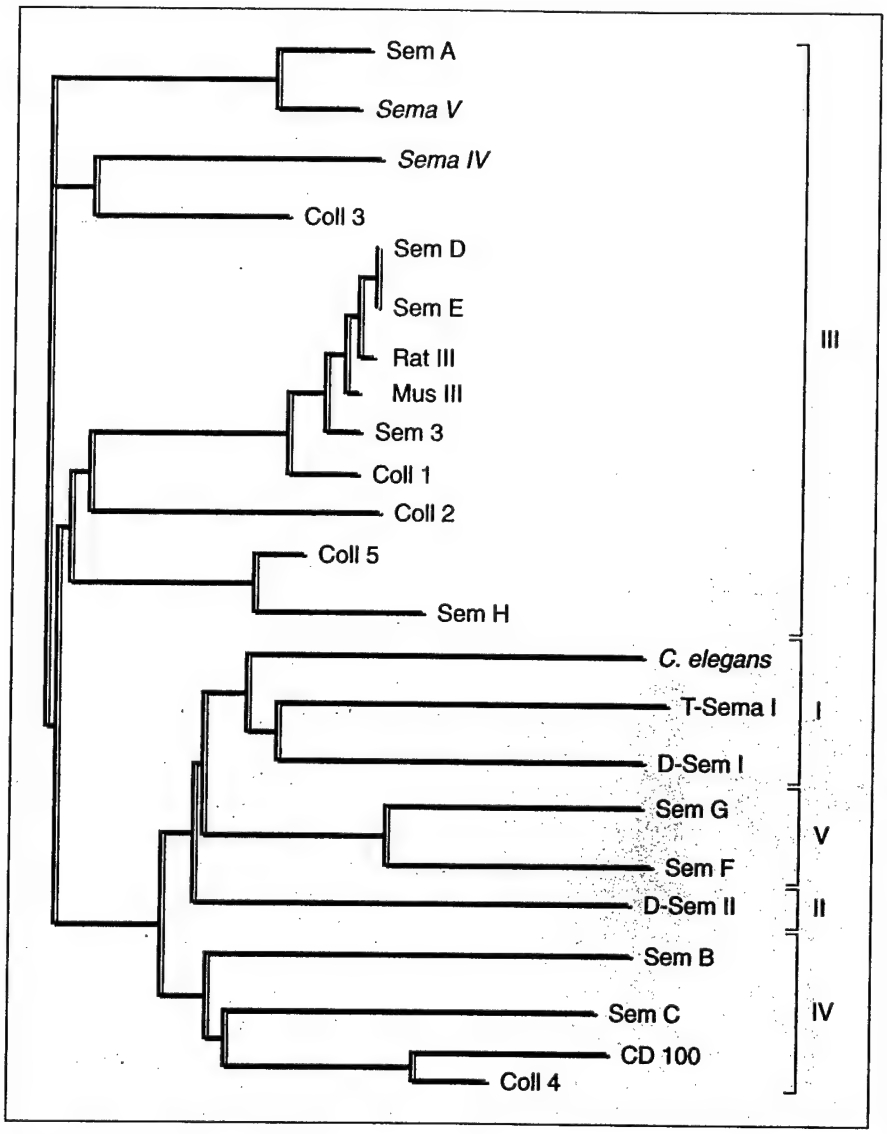


Figure 4. ***Arbre phylogénétique des sémaphorines.*** *Les séquences protéiques ont été obtenues par Genbank et l'arbre phylogénétique a été construit selon la méthode de Saitou et Nei [46]. Les chiffres romains sur la droite de la figure indiquent les différentes familles de sémaphorines.*

Quel est le mode de fonctionnement des sémaphorines?

On en connaît peu de choses pour l'instant, mais l'existence de deux groupes de protéines sécrétées et transmembranaires suggère qu'elles fonctionneraient, soit localement, soit à distance. Dans le cas des sémaphorines sécrétées, le domaine carboxy-terminal basique leur permettrait d'établir des liaisons avec des charges négatives sur la matrice extracellulaire ou sur la membrane cellulaire. Le domaine immunoglobuline pourrait contribuer à l'établissement d'interactions avec d'autres protéines. Les récepteurs de sémaphorines sécrétées ont été récemment clonés et identifiés: il s'agit de la neuropiline-1 et de la neuropiline-2 [38-41].

Dans le cas de la collapsine I (Coll I) de poulet, l'homologue de SemaIV, on a isolé un médiateur intracellulaire, le CRMP-62 *(collapsin response mediator protein)* [42]. Il est l'analogue de la protéine neuronale UNC-33 de nématode et il serait un intermédiaire entre Coll I et une protéine G; la collapsine activerait une cascade de transmission du signal par l'intermédiaire de protéines G trimériques. Cependant, le récepteur de Coll I reste pour l'instant inconnu [38-41].

Par ailleurs, on a identifié un récepteur des nétrines, autre classe de protéines impliquées dans le guidage des neurones [32, 33]. Il s'agit d'une protéine transmembranaire de la superfamille des immunoglobulines codée par le gène *DCC (deleted in colorectal colon)* [43]. Bien que la fonction de suppresseur de *DCC* vienne d'être remise en cause par des expériences d'invalidation de ce gène chez la souris *(m/s n° 2, vol. 13, p. 245)* [44], il est toutefois intéressant de noter que deux gènes, *DCC* et *Sema*, tous deux impliqués dans le contrôle du développement des cônes de croissance des neurones, sont localisés dans des régions chromosomiques soumises à des délétions dans deux types différents de tumeurs.

Les sémaphorines dans les cellules tumorales

Des études de séquençage d'ADNc obtenus par RT-PCR à partir de 28 lignées SCLC n'ont mis en évidence que des polymorphismes de l'ADN des gènes *SemaIV* et *SemaV* [21]. L'analyse de 120 lignées tumorales n'a pas montré de réarrangements des gènes codant pour les sémaphorines [22] sauf dans le cas des trois lignées SCLC décrites qui présentent une délétion homozygote des gènes *SemaIV* et *SemaV*; le gène *SemaIV* est exprimé dans un plus grand nombre de lignées SCLC étudiées que le gène *SemaV* [22]. Cependant, l'expression de *SemaIV* présente des variations, d'une part, au niveau de la quantité d'ARN messager et, d'autre part, au niveau du rapport des deux formes d'ARNm obtenues par épissage alternatif [16]. La signification de ces variations n'est pas comprise pour l'instant.

Les gènes *SemaIV*, *SemaV* et *GNAI2* sont localisés dans une région soumise à des pertes d'hétérozygotie et à des délétions homozygotes et ils sont donc des gènes candidats intéressants dans l'étude du développement des tumeurs. En outre, les gènes *SemaIV* et *GNAI2* sont inclus dans la séquence de 80 kb ayant un effet suppresseur de croissance des tumeurs. Enfin, l'hypothèse selon laquelle les sémaphorines activeraient une cas-

cade de transmission du signal mettant en jeu des protéines G trimériques ainsi que l'ensemble de leurs fonctions supposées renforcent l'intérêt de l'étude de ces gènes.

Conclusion

La présence d'un gène suppresseur a été clairement démontrée pour la région 3p25. En revanche, la situation est moins claire pour la région 3p14 dont les délétions résulteraient de l'instabilité chromosomique. De même, la participation de la région 3p21.3 dans le développement du cancer du poumon est complexe dans la mesure où des délétions dans deux zones différentes pourraient avoir des effets indépendants, cumulatifs ou synergiques; la participation de plusieurs gènes, présents dans cette région, dans le développement des tumeurs mériterait d'être évaluée ■

Remerciements

Les auteurs remercient l'ARC (Association pour la Recherche sur le Cancer) pour son soutien financier des travaux concernant l'étude des gènes de la région 3p21.3 en cours de réalisation dans l'ESA Cnrs 6031 (IBMIG, Université de Poitiers). Les travaux effectués à l'Université du Colorado ont bénéficié du soutien financier du NIH (R01 CA68383). La collaboration entre les deux laboratoires est aidée par une « Subvention à la Recherche Collaborative » de la part de l'OTAN (Organisation du Traité de l'Atlantique Nord).

RÉFÉRENCES

1. Martinet Y, Brambilla E, Martin JP, Martinet N, Vignaud JM. Les bases biologiques de la prévention du cancer du poumon. *Med Sci* 1997; 13: 1465-71.

2. Kok K, Naylor S, Buys C. Deletions of the short arm of chromosome 3 in solid tumors and the search for suppressor genes. *Adv Cancer Res* 1997; 71: 27-92.

3. Rabbitts P, Douglas J, Daly M, Sundaresan V, Fox B, Haselton P, Wells F, Albertson D, Waters J, Bergh J. Frequency and extent of allelic loss in the short arm of chromosome 3 in non small-cell lung cancer. *Genes Chromosom Cancer* 1989; 1: 95-105.

4. Todd S, Franklin W, Varella-Garcia L, Kennedy T, Hilliker K, Hahner L, Anderson M, Wiest J, Drabkin H, Gemmill B. Homozygous deletions of human chromosome 3p in lung tumors. *Cancer Res* 1997; 57: 1344-52.

5. Latif F, Tory K, Gnarra J, Yao M, Duh FM, *et al.* Identification of the von Hippel-Lindau disease tumor suppressor gene. *Science* 1993; 260: 1317-20.

6. Duan DR, Pause A, Burgess WH, Aso T, Chen DY, Garrett KP, Conaway RC, Conaway JW, Linehan WM, Klausner RD. Inhibition of transcription elongation by the VHL tumor suppressor protein. *Science* 1995; 269: 1402-6.

7. Killary AM, Wolf ME, Giambernardi TA, Naylor SL. Definition of a tumor suppressor locus within human chromosome 3p21-p22. *Proc Natl Acad Sci USA* 1992; 89: 10877-81.

8. Sanchez Y, el-Naggar A, Pathak S, Killary AM. A tumor suppressor locus within 3p14-p12 mediates rapid cell death of renal cell carcinoma *in vivo*. *Proc Natl Acad Sci USA* 1994; 91: 3383-7.

9. Cohen AJ, Li FP, Berg S, Marchetto DJ, Tsai S, Jacobs SC, Brown RS. Hereditary renal-cell carcinoma associated with a chromosomal translocation. *N Engl J Med* 1979; 301: 592-5.

10. Boldog FL, Gemmill RM, Wilke CM, Glover TW, Nilsson AS, Chandrasekharappa SC, Brown RS, Li FP, Drabkin HA. Positional cloning of the hereditary renal carcinoma 3;8 chromosome translocation breakpoint. *Proc Natl Acad Sci USA* 1993; 90: 8509-13.

11. Glover TW, Berger C, Coyle J, Echo B. DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Hum Genet* 1984; 67: 136-42.

12. Boldog F, Gemmill R, West J, Robinson M, Robinson L, *et al.* 3p14 homozygous deletions and sequence analysis of FRA3B. *Hum Mol Genet* 1997; 6: 193-203.

13. Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, *et al.* The *FHIT* gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. *Cell* 1996; 84: 587-97.

14. Daly MC, Xiang RH, Buchhagen D, Hensel CH, Garcia DK, Killary AM, Minna JD, Naylor SL. A homozygous deletion on chromosome 3 in a small cell lung cancer cell line correlates with a region of tumor suppressor activity. *Oncogene* 1993; 8: 1721-9.

15. Kok K, van den Berg A, Veldhuis PM, van der Veen AY, Franke M, *et al.* A homozygous deletion in a small cell lung cancer cell line involving a 3p21 region with marked instability in yeast artificial chromosomes. *Cancer Res* 1994; 54: 4183-7.

16. Roche J, Boldog F, Robinson M, Robinson L, Varella-Garcia M, Swanton M, Waggoner B, Fishel R, Franklin W, Gemmill R, Drabkin H. Distinct 3p21.3 deletions in lung cancer, analysis of deleted genes and identification of a new human semaphorin. *Oncogene* 1996; 12: 1289-97.

17. Magovcevic I, Ang SL, Seidman JG, Tolman CJ, Neer EJ, Morton CC. Regional localization of the human G protein alpha i2 (GNAI2) gene: assignment to 3p21 and a related sequence (GNAI2L) to 12p12-p13. *Genomics* 1992; 12: 125-9.

18. Murata Y, Tamari M, Takahashi T, Horio Y, Hibi K, Yokoyama S, Inazawa J, Yamakawa K, Ogawa A. Characterization of an 800 kb region at 3p22-p21.3 that was homozygously deleted in a lung cancer cell line. *Hum Mol Genet* 1994; 3: 1341-4.

19. Papadopoulos N, Nicolaides NC, Wei YF, Ruben SM, Carter KC, *et al.* Mutation of a *mutL* homolog in hereditary colon cancer. *Science* 1994; 263: 1625-9.

20. Rudolph U, Finegold MJ, Rich SS, Harriman GR, Srinivasan Y, Brabet P, Boulay G, Bradley A, Birnbaumer L. Ulcerative colitis and adenocarcinoma of the colon in G alpha i2-deficient mice. *Nat Genet* 1995; 10: 143-8.

21. Xiang RH, Hensel CH, Garcia DK, Carlson HC, Kok K, Daly MC, Kerbacher K, van den Berg A, Veldhuis P, Buys C, Naylor SL. Isolation of the human semaphorin III/F gene *(SEMA3F)* at chromosome 3p21, a region deleted in lung cancer. *Genomics* 1996; 32: 39-48.

22. Sekido Y, Bader S, Latif F, Chen JY, Duh FM, Wei MH, Albanes JP, Lee CC, Lerman MI, Minna JD. Human semaphorins A (V) and (IV) reside in the 3p21.3 small cell lung cancer deletion region and demonstrate distinct expression patterns. *Proc Natl Acad Sci USA* 1996; 93: 4120-5.

23. Wei MH, Latif F, Bader S, Kashuba V, Chen JY, *et al.* Construction of a 600-kilobases clone contig and generation of a transcriptional map surrounding the lung cancer tumor suppressor gene *(TSG)* locus on human chromosome 3p21.3: progress toward the isolation of a lung cancer TSG. *Cancer Res* 1996; 56: 1487-92.

24. Todd MC, Xiang RH, Garcia DK, Kerbacher KE, Moore SL, *et al.* An 80 kb P1 clone from chromosome 3p21.3 suppresses tumor growth *in vivo*. *Oncogene* 1996; 13: 2387-96.

25. Kolodkin AL, Matthes DJ, Goodman CS. The semaphorin genes encode a family of transmembrane and secreted growth cone guidance molecules. *Cell* 1993; 75: 1389-99.

26. Kolodkin AL. Semaphorins: mediators of repulsive growth cone guidance. *Trends Cell Biol* 1996; 6: 15-22.

27. Püschel AW. The semaphorins: a family of axonal guidance molecule. *Eur J Neurosc* 1996; 8: 1317-21.

28. Ensser A, Fleckenstein B. Alcelaphine herpesvirus type 1 has a semaphorin-like gene. *J Gen Virol* 1995; 76: 1063-7.

29. Püschel AW, Adams R, Betz H. Murine semaphorine D/collapsin is a member of a diverse gene family and creates domains inhibitory for axonal extension. *Neuron* 1995; 14: 941-8.

30. Messersmith E, Leonardo E, Shatz C, Tessier-Lavigne M, Goodman C, Kolodkin A. Semaphorin III can function as a selective chemorepellent to pattern sensory projections in the spinal cord. *Neuron* 1995; 14: 949-59.

31. Keynes R, Cook G. Axon guidance molecules. *Cell* 1995; 83: 161-9.

32. Bloch-Gallego E, Sotelo C. Chimioattraction ou chimiorépulsion axonale: rôle des nétrines et des sémaphorines. *Med Sci* 1998; 14: 44-52.

RÉFÉRENCES

33. Tessier-Lavigne M, Goodman C. The molecular biology of axone guidance. *Science* 1996; 274: 1123-33.

34. Behar O, Golden JA, Mashimo H, Schoen FJ, Fishman MC. Semaphorin III is needed for normal patterning and growth of nerves, bones and heart. *Nature* 1996; 383: 525-8.

35. Adams RH, Betz H, Püschel AW. A novel class of murine semaphorins with homology to thrombospondin is differentially expressed during early embryogenesis. *Mech Dev* 1996; 57: 33-45.

36. Hall K, Boumsell L, Schultze J, Boussiotis V, Dorfman D, Cardoso A, Bensussan A, Nadler L, Freeman G. Human CD100, a novel leukocyte semaphorin that promotes B-cell aggregation and differentiation. *Proc Natl Acad Sci USA* 1996; 93: 11780-5.

37. Yamada T, Endo R, Gotoh M, Hirohashi S. Identification of semaphorin E as non-MDR drug resistance gene of human cancers. *Proc Natl Acad Sci USA* 1997; 94: 14713-8.

38. He Z, Tessier-Lavigne M. Neuropilin is a receptor for the axonal chemorepellent semaphorin III. *Cell* 1997; 90: 739-51.

39. Kolodkin A, Levengood D, Rowe E, Tai Y, Giger R, Ginty D. Neuropilin is a semaphorin III receptor. *Cell* 1997; 90: 753-62.

40. Feiner L, Koppel AM, Kobayashi H, Raper JA. Secreted chick semaphorins bind recombinant neuropilin with similar affinity but bind different subsets of neurons *in situ*. *Neuron* 1997; 19: 539-45.

41. Chen H, Chédotal A, He Z, Goodman CS, Tessier-Lavigne M. Neuropilin-2 a novel member of the neuropilin family is a high affinity receptor for the semaphorins Sema E and Sema IV but not Sema III. *Neuron* 1997; 19: 547-59.

42. Goshima Y, Nakamura F, Strittmatter P, Strittmatter S. Collapsin-induced growth cone collapse mediated by an intracellular protein related to UNC-33. *Nature* 1995; 376: 509-14.

43. Keino-Massu K, Masu M, Hinck L, Leonardo E, Chan S, Culotti J, Tessier-Lavigne M. *Deleted in Colorectal Cancer (DCC)* encodes a netrin receptor. *Cell* 1996; 87: 175-85.

44. Fazeli A, Dickinson SL, Hermiston ML, Tighe RV, Steen RG, *et al.* Phenotype of mice lacking functional *Deleted in colorectal cancer (Dcc)* gene. *Nature* 1997; 386: 796-804.

45. Rabbitts P, Bergh J, Douglas J, Collins F, Waters J. A submicroscopic homozygous deletion at the D3S3 locus in a cell line isolated from a small cell lung carcinoma. *Genes Chromosom Cancer* 1990; 2: 231-8.

46. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987; 4: 406-25.

TIRÉS À PART

J. Roche.

Summary

Chromosome 3p, tumor suppressors and new semaphorin genes at 3p21.3

Loss of chromosome 3p material is a critical event in the pathogenesis of lung cancer and other carcinomas. Loss of heterozygosity (LOH) studies have defined at least three different regions (3p25, 3p21.3 and 3p14-cen) that may contain tumor suppressor genes *(TSG)*. The identification of homozygous deletions involving 3p12-13, 3p14 and 3p21.3 in cell lines and direct tumors support the LOH results and have been used to further define the important segments. The *VHL (von Hippel Lindau)* gene from 3p25 is a tumor suppressor but is rarely involved in non-syndrome tumors including lung cancer. The 3p14.2 region contains the fragile site, FRA3B, a hereditary renal carcinoma-associated 3;8 translocation and a putative tumor suppressor gene, *FHIT*. However, our studies indicate that *FHIT* is not selectively lost and therefore may not be a *TSG*. In contrast to the selective loss of a tumor suppressor gene, an alternative hypothesis is that some recurrently deleted regions such as FRA3B could undergo loss as a consequence of genomic instability. Distinct homozygous deletions have been identified in 3p21.3: the more distal is adjacent to the DNA mismatch repair gene *hMLH1* and the more proximal includes *GNAI2* coding for the Gαi2 subunit of a G protein, and two genes coding for Semaphorins IV and V. Semaphorins/collapsins were first identified as nerve growth cone guidance signals that act as selective repulsive factors. However, their wide expression suggests that they could serve other functions such as cell differentiation or recognition. As a consequence, their inactivation could be potentially important in tumor development.

Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 9572–9577, August 1998
Medical Sciences

The hereditary renal cell carcinoma 3;8 translocation fuses *FHIT* to a *patched*-related gene, *TRC8*

ROBERT M. GEMMILL*†, JAMES D. WEST*, FERENC BOLDOG*‡, NAOTAKE TANAKA*, LINDA J. ROBINSON*, DAVID I. SMITH§, FREDERICK LI¶, AND HARRY A. DRABKIN*

*Division of Medical Oncology, University of Colorado Health Sciences Center, 4200 East 9th Avenue, Denver, CO 80262; §Department of Laboratory Medicine and Pathology, Mayo Foundation, Rochester, MN 55905; and §Harvard School of Public Health, 44 Binney Street, Boston, MA 02115

Communicated by David Marshall Prescott, University of Colorado, Boulder, CO, June 8, 1998 (received for review May 15, 1998)

ABSTRACT **The 3;8 chromosomal translocation, t(3;8)(p14.2;q24.1), was described in a family with classical features of hereditary renal cell carcinoma. Previous studies demonstrated that the 3p14.2 breakpoint interrupts the fragile histidine triad gene (*FHIT*) in its 5′ noncoding region. However, evidence that *FHIT* is causally related to renal or other malignancies is controversial. We now show that the 8q24.1 breakpoint region encodes a 664-aa multiple membrane spanning protein, TRC8, with similarity to the hereditary basal cell carcinoma/segment polarity gene, *patched*. This similarity involves two regions of *patched*, the putative sterol-sensing domain and the second extracellular loop that participates in the binding of sonic hedgehog. In the 3;8 translocation, TRC8 is fused to *FHIT* and is disrupted within the sterol-sensing domain. In contrast, the *FHIT* coding region is maintained and expressed. In a series of sporadic renal carcinomas, an acquired *TRC8* mutation was identified. By analogy to *patched*, *TRC8* might function as a signaling receptor and other pathway members, to be defined, are mutation candidates in malignant diseases involving the kidney and thyroid.**

The 3;8 chromosomal translocation, t(3;8)(p14.2;q24.1), was described in a family with classical features of hereditary renal cell carcinoma (RCC), i.e., autosomal dominant inheritance, and early onset and bilateral disease (1). The translocation and RCC segregated concordantly, and a follow-up analysis reported the occurrence of thyroid cancer in two translocation carriers with RCC (2). Frequent 3p loss of heterozygosity in sporadic RCC led to the initial assumption that a critical tumor suppressor gene (TSG) would be located at 3p14. However, identification of the von Hippel-Lindau (*VHL*) gene at 3p25 (3), which frequently is mutated in RCC, provided an alternative explanation for at least some observed 3p loss of heterozygosity. Also, van den Berg and Buys (4) subsequently have reported that region 3p21 may be a target involved in the malignant progression of renal tumors.

Within 3p14, Ohta *et al.* (5) identified the fragile histidine triad gene (*FHIT*), which was interrupted in its 5′ untranslated region by the 3;8 translocation. The human gene, like its yeast homologue, encodes diadenosine-5′,5‴-P^1,P^3-triphosphate hydrolase activity (6), an unprecedented TSG function. Although several reports have described *FHIT* alterations in diverse carcinomas by using nested reverse transcriptase–PCR (RT-PCR) (5, 7–10), other results have been contradictory (11–17). In fact, most *FHIT* abnormalities occur in the presence of wild-type transcripts and result from low-abundance splicing alterations, similar to those seen for *TSG101* (13, 18). We described a series of 3p14 homozygous deletions, primarily in cervical and colorectal carcinoma cell lines, which coincided with FRA3B, the most inducible fragile region in the genome (14). Interestingly, p53 alterations appeared to be a prerequisite. The proximity of *FHIT* exon 5 with FRA3B suggested that its loss might be primarily related to genomic instability in contrast to negative selection during tumor development (14).

These results made *FHIT* an unlikely, or at least suspect, causative gene in the hereditary t(3;8) family and led us to continue a search for alternatives. We noted that *FHIT*, in one parotid adenoma, underwent fusion with the high mobility group protein gene (*HMGIC*), the causative gene in a variety of benign tumors (19). In fact, other *HMGIC* translocations with unrelated genes indicated that *FHIT* could be a bystander in this fusion. However, it suggested that the 3;8 translocation might fuse *FHIT* to an alternative candidate gene on chromosome 8. By using 5′ rapid amplification of cDNA ends (RACE) (20), we were able to identify a gene, *TRC8*, which was related to the *Drosophila* segment polarity gene *patched*. In addition to being the receptor for sonic hedgehog (21, 22), *patched* is responsible for both hereditary and sporadic basal cell carcinomas as well as medulloblastomas (23–25). Together with the identification of a *TRC8* mutation in a sporadic renal carcinoma, our results indicate that *TRC8* may define an additional pathway of mutations leading to the development of renal and thyroid cancer.

MATERIALS AND METHODS

Tumors, Cell Lines, and Genomic Clones. Tumor cell lines were obtained from the American Type Culture Collection, and somatic cell hybrids were generated by us previously (26). The hybrids TL12–8 and 3;8/4–1 contain the der(3) and der(8) chromosomes, respectively, from the t(3;8) lymphoblastoid cell line TL9944 (without either normal 3 or 8 chromosomes). The human lymphoblastoid line AG4103 served as a normal control. Isolation of DNA and RNA was performed by using standard methods. The HD-7 genomic phage clone carrying the 3;8 translocation breakpoint from the der(8) chromosome was isolated from a library prepared from the TL9944 cell line in λ FIXII (Stratagene). A chromosome 3 probe (λ4040; ref. 27), which maps just distal to the 3;8 breakpoint and detects the rearrangement, was used for screening this library.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: RCC, renal cell carcinoma; TSG, tumor suppressor gene; *FHIT*, fragile histidine triad gene; RACE, rapid amplification of cDNA ends; *TRC8*, translocation in renal carcinoma, chromosome 8 gene; TM, transmembrane segment; SSD, sterol-sensing domain; RT-PCR, reverse transcriptase–PCR; SSCP, single-stranded conformational polymorphism analysis; HMG, hydroxymethylglutaryl; YAC, yeast artificial chromosome.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF064800 and AF064801).

†To whom reprint requests should be addressed. e-mail: gemmill@loki.uchsc.edu.
‡Present address: CuraGen Corporation, 12085 Research Drive, Alachua, FL 32615.

RACE. 5′ RACE was performed essentially as described (28). First-strand cDNA synthesis was primed by using a *FHIT* exon 4-specific primer R1 (TCAGAAGACTGCTACCTCTTCG), followed by dCTP tailing with terminal deoxynucleotidyl transferase. Primary amplification used the AAP 5′ RACE primer together with a nested *FHIT* exon 4-specific primer R2 (TCAGTGGCAGGATGCACAG). Second-round PCR used primer AUAP with a second nested *FHIT* exon 4-specific primer R3 (GGTCTAAGCAGGCAGGTATTC). Products were cloned into a T-vector (pBluescript II KS), analyzed by hybridization with additional internal *FHIT* oligonucleotides F4 (TGGAAGGGAGAGAAAGAG) and R4 (GGTATTCCTAGGATAC), and sequenced.

Chromosomal Localization and RT-PCR Analysis. PCR mapping used *TRC8*-specific primers R-M (GCCCTGCCTTTACATCATCGAC) and F-O (AGATCTGGAGCACGATGCAGAAC). PCRs were performed under touch-down annealing conditions with Perkin–Elmer Buffer II and Promega AmpliTaq DNA polymerase. Touch-down annealing temperatures started at 70°C and ended at 60°C (ΔT of −0.5°C per cycle) for 20 cycles, followed by 15 cycles at 60°C. Products generated from 10 to 40 ng of template were separated on a 2.0% agarose gel. cDNA synthesis used random hexamer primers along with Superscript II (Life Technologies, Gaithersburg, MD). Subsequent PCRs were performed as above, except touch-down annealing temperatures were adjusted to 65°C–55°C. The EMR primer, specific for the 3′ portion of *TRC8*, was TCTTGTTAGCCAAAAGACTCG, whereas the F1 primer specific for *FHIT* exon 1 was TCCCTCTGCCTTTCATTCC.

DNA Sequence Analysis. Sequencing was performed on an ABI377 through the Colorado Cancer Center DNA Sequencing Core. Analysis for transmembrane segments (TMs) was performed by using five prediction programs: PHD_htm at EMBL (http://www.embl-heidelberg.de/predictprotein/), TMpred at ISREC (http://ulrec3.unil.ch/software/TMPRED_form.html), SOSUI at Tokyo University (http://www.tuat.ac.jp/~mitaku/adv_sosui/), DAS at Stockholm University (http://www.biokemi.su.se/~server/DAS/), and PSORT at Osaka University (http://psort.nibb.ac.jp/). All 10 TMs that are underlined in Fig. 1*A* were predicted by at least four of the five programs, although in most cases the programs did not agree on the precise boundaries of the segment.

Single-Stranded Conformational Polymorphism Analysis (SSCP). SSCP was performed by the method of Spritz *et al.* (29). Nine primer sets were designed to amplify the entire coding region in segments averaging 325 bp and that also would span any intron-exon boundaries. These primer sets were: set 1 (set1F AGTTGCCCGCCTTAGCC and set1R CCAAAGACACATACTCGACCC), set 2 (set2F CATAACTCTTAGTGGGGAAACATTC and set2R TGTAACGTATCCAATTCCAAATG), set 3 (set3F TGGCACTTATCGTTCTACAGC and set3R TCTTGTTAGCCAAAAGACTCG), set 4 (set4F AGTGTTTGTCCTGGCAGTG and set4R ACAGTTAGTGTAGAATCGCACCC), set 5 (set5F TGGCAAATGAAACTGATTCC and set5R CATGGATAAAATGCAGGACTG), set 6 (set6F AAGACCAGAAGAGAGACTTATTCG and set6R TGCTGTAACTGCAAACAACC), set 7 (set7F TCTTTGGCATCACTATGCAC and set7R CTTCACAGCAGTCCTACGATTC), set 8 (set8F CCAAAAATGGCTGGAAGAC and set8R TGTCAGATTCAGCAGCAGC), and set 9 (set9F CCACCCAATGAAACTCCAG and set9R AGTAGCACATCACAGTAAACGG). In addition, three primer sets were designed to amplify the 5′ untranslated region including: set P1 (setP1F TCCCAGGCAGCTCTGAAC and setP1R ACCATCTTGACCTCGCCC), set P2 (setP2F GTTCGCTTGACTGACGGC and setP2R ATGAGCCGCTGCCACAC), and set P3 (setP3F CACCGAAACCCAGAGACC and setP3R CCAAAGACACATACTCGACCC). These primers were used to amplify genomic DNA (10 ng) under 65°C to 55°C touch-down conditions, as described above. Because of the

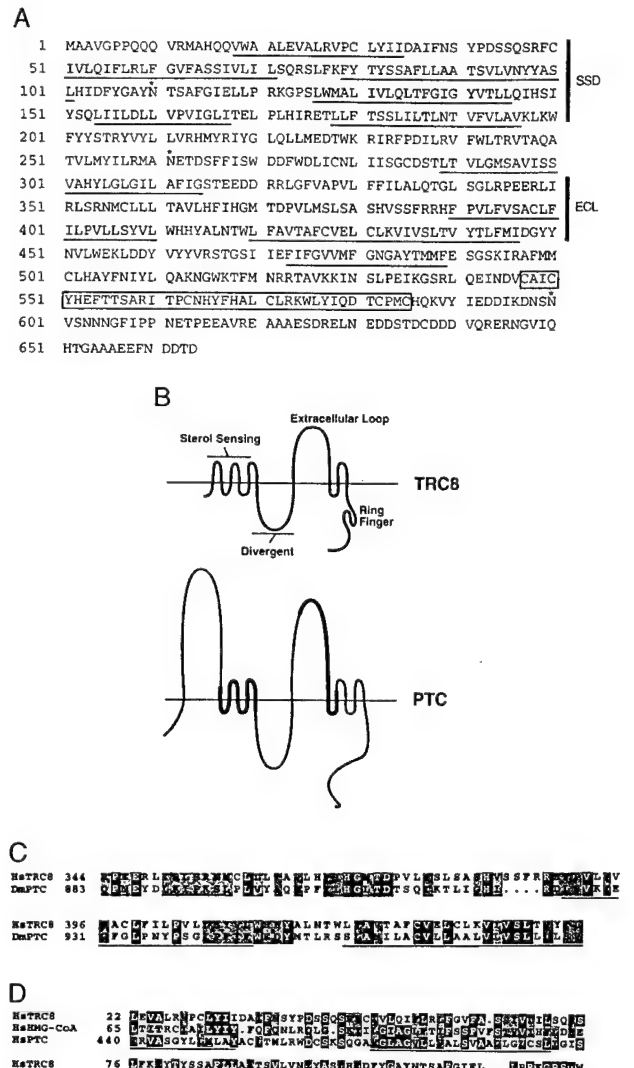


FIG. 1. (*A*) Predicted amino acid sequence of *TRC8*. The sequence begins with the first methionine present in the isolated cDNAs. The 3;8 translocation breakpoint occurs between amino acids 60 and 61 (arrowhead). Predicted TMs are underlined, and three potential N-linked glycosylation sites are indicated by *. Two regions showing similarity to *patched* [the SSD and the extracellular loop (ECL)] are shaded. The RING-H2 motif is boxed. (*B*) Schematic of the predicted structure for *TRC8* compared with *patched* (PTC). The horizontal line represents the lipid bilayer that is crossed by 10 putative TMs of *TRC8* and 12 in *patched*. The divergent loop refers to a region of *patched* that is nonconserved among patched homologues (45). The diagram of *patched* shows regions that are similar or different from *TRC8* by bold black or thin shaded lines, respectively. The N-terminal extracellular loop of *patched* is absent from *TRC8*. The predicted intracellular loop 3′ of the SSD is not conserved between either of the two known murine *patched* genes (*Ptch1* and *Ptch2*) or in *TRC8*. (C) Amino acid sequence homology between *TRC8* and Drosophila *patched*. The Dm *Ptc* sequence (GenBank accession no. M28999) was aligned with *TRC8* by gapped BLAST. Identical amino acids are indicated by white letters on black, and similar amino acids (positive scores in a PAM250 matrix) are shaded. Two TRC8 TMs within this homology region are underlined. (*D*) Alignment within the putative SSD. Human sequences for HMG-CoA reductase (Swissprot accession no. P04035) and *Patched* (GenBank accession no. U59464) were aligned with *TRC8* by gapped BLAST. Identities and similarities are indicated as in *C*; TMs within the putative SSD of *TRC8* are underlined.

high GC content present in a 300-bp segment near the amino terminus of TRC8, PCRs for primer sets 1, P2, and P3 contained 2.5 M betaine (30). Reaction products were mixed 50:50 with denaturing dyes (95% formamide, 10 mM NaOH, 20 mM EDTA, 0.02% bromophenol blue, and 0.02% xylene cyanol) and heated to 95°C for 5 min immediately before loading. Samples were separated at 8 W for 16 hr on 0.5× MDE (FMC) gels containing 0.6× Tris-borate buffer and 10% glycerol. Bands were visualized by silver staining.

RESULTS

Identification and Sequence Analysis of *TRC8*. To test for a possible fusion gene, RNA was isolated from TL9944 lymphoblastoid cells carrying the 3;8 translocation and subjected to RACE. The t(3;8) breakpoint interrupts *FHIT* between exons 3 and 4. Therefore, 5′ RACE products were generated by using nested primers within exon 4. Although the cloned amplification products were identified by hybridization to an internal *FHIT* exon 4 oligonucleotide (R4), nearly 80% of these clones were negative for a *FHIT* exon 3 oligonucleotide (F4) suggesting they might represent a gene fusion. Twelve such clones were further examined by DNA sequencing, and nine were found to contain an identical novel sequence spliced exactly to the 5′ end of *FHIT* exon 4. Mapping studies confirmed that the sequences were derived from chromosome 8 (see below). We refer to this gene as *TRC8* for translocation in renal cancer from chromosome 8.

The coding region of *TRC8* was determined from multiple cDNA clones and PCR products isolated from a human fetal brain library (Stratagene) plus IMAGE clone 331H8 identified from dbEST. Additional genomic sequences extending approximately 2.1 kb upstream of the first in-frame methionine were obtained from phage HD7, which contains both chromosome 3 and 8 material spanning the 3;8 breakpoint. The cDNA sequence contains a predicted 1,992-bp ORF encoding 664 amino acids (Fig. 1*A*). The upstream cDNA and genomic sequences are GC-rich, indicative of a CpG island. The 5′ most cDNA clone extended to position −286 bp, and additional RACE experiments failed to extend this further. Using a promoter prediction program (http://www-hgc.lbl.gov/projects/promoter.html), four transcriptional start sites were located at −622, −55, −36, and −22 bp of the first methionine. The −22 site corresponded precisely to the two longest RACE products. Thus, it appears that multiple transcription start sites are used.

The ORF is predicted to encode a 664-aa protein of 76 kDa with 10 membrane-spanning segments (Fig. 1*B*). *TRC8* contains two regions of similarity with *patched,* the receptor for sonic hedgehog (SHH) (21, 22). The region from amino acids 344–443 shows the strongest match with 60% similarity and 23% identity (Fig. 1*C*). The corresponding segment of *patched* spans amino acids 883–979 and represents most of the second predicted extracellular domain involved in the binding of SHH (21). A second region of *patched* similarity involves amino acids 22–179 and encodes a putative sterol-sensing domain (SSD), Fig. 1D. Such domains have been identified in hydroxymethyl glutaryl (HMG)-CoA reductase and the sterol regulatory element binding protein (SREBP) cleavage activating protein (SCAP) (31). *Patched* contains a putative SSD, of unknown function, from amino acids 440 to 600 (32, 33). This region is 53% similar/17% identical to the SSD of HMG-CoA reductase (34); the corresponding region from *TRC8* shows 63% similarity/17% identity. We surmise, based on the multiple TMs and regions of *patched* similarity, that *TRC8* encodes a membrane-bound receptor.

In addition, a perfect match with a ring-finger motif of the RING-H2 subtype (35) was found between amino acids 547 and 585. The RING-H2 motif ($CX_2CX_{9-27}CXHX_2HX_2CX_{6-17}CX_2C$) differs from the standard RING finger by replacement of the fourth cysteine with a histidine. Functionally, RING-H2 motifs have been suggested to be protein–protein or protein–lipid interaction domains (35). We also observed that *TRC8* is highly conserved among mammals. This finding was evident from two murine expressed sequence tags (dbEST clones mu78h12 and vl43c01) found to be 93% and 89% identical, respectively, at the nucleotide level over 971 bp.

Expression of *TRC8*. Hybridization of *TRC8* to a Northern blot (CLONTECH) prepared from adult human tissues and placenta identified a message of approximately 3.0 kb (Fig. 2*A*). Although our longest cDNA clones total 2.5 kb, use of the −622-bp promoter, as discussed above, would result in a 2.9-kb message, close to the observed size. Although expression in the lung and kidney appeared reduced, hybridization with a control glyceraldehyde-3-phosphate dehydrogenase probe (data not shown) indicated that there was less RNA present in these samples. A human RNA dot blot (RNA Master blot, CLONTECH) revealed *TRC8* message in all tissues examined (Fig. 2*B*) with the highest levels in testis (D1), placenta (F4), and adrenal (D5). *TRC8* is expressed in both fetal (G3) and adult kidney (E1) and in adult thyroid (D6), the suspected target organs for *TRC8* aberrations in the 3;8 translocation family.

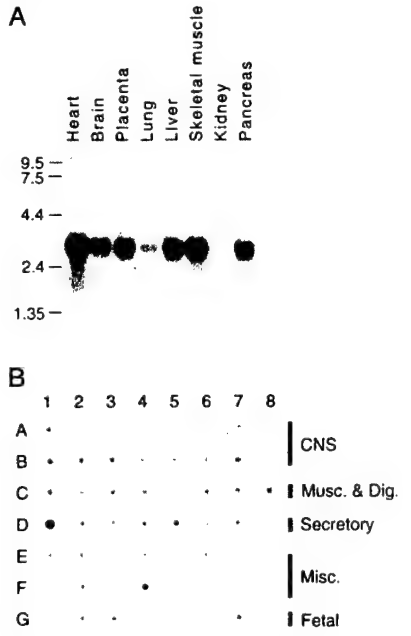


FIG. 2. Analysis of *TRC8* expression by Northern (*A*) and dot blot (*B*). (*A*) Gel resolved polyadenylated RNA (2 μg) from adult human tissues (CLONTECH) was hybridized under recommended conditions with a 1.5-kb 3′ *TRC8* cDNA encompassing most of the TMs and the ring finger (bp 83–1623). A second, largely nonoverlapping probe (bp 1446–2212) yielded essentially the same pattern. The filter was exposed for 18 hr at −80°C. (*B*) A CLONTECH human RNA master dot blot was hybridized with the same probe as in *A* under recommended conditions and exposed for 15 hr. Final wash conditions were 0.1× standard saline citrate, 0.5% SDS at 55°C for 20 min. Signals were collected on a Molecular Dynamics PhosphorImager. Blank positions included B8, F5-F8, and G8. Central nervous system tissues (A1-A8 and B1-B7) included (in order) whole brain, amygdala, caudate nucleus, cerebellum, cerebral cortex, frontal lobe, hippocampus, medulla oblongata, occipital lobe, putamen, substantia nigra, temporal lobe, thalamus, sub-thalamic nucleus, and spinal cord. Musculature and digestive tissues (C1-C8) included heart, aorta, skeletal muscle, colon, bladder, uterus, prostate, and stomach. Secretory tissues (D1-D8) included testis, ovary, pancreas, pituitary, adrenal, thyroid, salivary, and mammary glands. Miscellaneous tissues (E1-E8 and F1-F4) included kidney, liver, small intestine, spleen, thymus, peripheral leukocytes, lymph node, bone marrow, appendix, lung, trachea, and placenta. Fetal tissues (G1-G7) included brain, heart, kidney, liver, spleen, thymus, and lung. All control spots (yeast and *Escherichia coli* RNAs, human Cot1, and total human DNAs) were blank (not shown).

Mapping of *TRC8*. *TRC8* sequences were localized to the immediate region of the breakpoint on chromosome 8 by both PCR and Southern blot analysis of hybrids, yeast artificial chromosome (YACs), and phage clones (Fig. 3). Primers derived from the 5′ coding portion of *TRC8* yielded the expected product in the chromosome 8-only hybrid (Fig. 3*A*, lane 4) but not in a chromosome 3-only hybrid (lane 3). The same product was also present on the der(8) but not the der(3) chromosome from the 3;8 translocation. Similarly, the 8q24 YAC 880A9 was positive as was the lambda clone, HD7, which contained both chromosome 8 and 3 material from the breakpoint junction. Southern blot analysis (Fig. 3*B*) demonstrated that the remaining 3′ portion of TRC8 was contained on the der(3) chromosome (lane 5). Notably, the probe hybridized to an altered band (arrow) in the der(3) hybrid consistent with the t(3;8) rearrangement. Together, these data indicate that *TRC8* is localized to 8q24 and is interrupted by the 3;8 translocation and that its 5′ to 3′ orientation is centromere to telomere.

Both Reciprocal Products are Expressed in t(3;8) Lymphoblastoid Cells. To determine whether both reciprocal products were expressed, we performed RT-PCR analysis on RNA isolated from TL9944 lymphoblastoid cells carrying the 3;8 translocation. Primers that flanked the breakpoint and were specific for the 5′ and 3′ portions of either *TRC8* or *FHIT* were used. As can be seen in Fig. 4, primers specific for wild-type *FHIT* and *TRC8* generated bands of the expected size from both t(3;8) and control RNAs. In contrast, the 5′ *TRC8* primer together with the 3′ *FHIT* primer produced a product only from the (3;8) translocation cell line. Similarly, the reciprocal 5′ *FHIT*/3′ *TRC8* product also was observed only in the translocation line. No products were detected in the absence of RT. Sequence analysis (Fig. 4*B*) confirmed that these unique RT-PCR products contained fusions of *TRC8* and *FHIT* sequences, as expected. Thus, although *FHIT* is interrupted in its 5′ untranslated region, its complete coding sequence is contained in the product from the der(8) chromosome. In contrast, *TRC8* is interrupted within the predicted SSD.

A Tumor-Specific Mutation in Sporadic RCCs. An SSCP analysis was performed by using 12 primer pairs covering the coding sequence and the 5′ untranslated region in 32 renal carcinomas. A duplication of 12 nucleotides in the 5′ untranslated region was identified in one tumor that was absent in matched normal DNA and is thus tumor specific (Fig. 5*A*, lane 5). This mutation was verified by multiple separate PCR amplifications, SSCP analyses and sequencing, as well as by the use of an alternative primer set, thus eliminating the possibility of a PCR artifact. The duplication (Fig. 5*B*) was found to occur in a consistent RNA stem loop structure predicted by the GCG program MFOLD in both energetically optimal and all suboptimal folds.

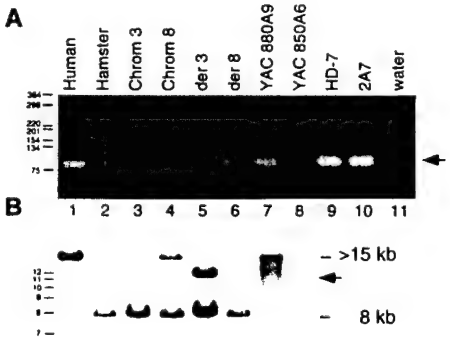


FIG. 3. (*A*) Localization of 5′ *TRC8* sequences to chromosome 8q. Primers R-M and F-O amplify an 82-bp fragment specific for the 5′ portion of *TRC8*. Templates in lanes 1–11 included, respectively, AG4103 (normal human), CHO glyA$^-$ (hamster), UCTP-2A3 (chromosome 3 only hybrid), 706-B6, clone17 (chromosome 8 only hybrid), TL12–8 [t(3;8) der(3) hybrid], 3;8/4–1 [t(3;8) der(8) hybrid], YAC 880A9 (chromosome 8-specific YAC spanning 3;8 breakpoint), YAC 850A6 (chromosome 3-specific YAC spanning 3;8 breakpoint), HD-7 (genomic phage clone carrying the 3;8 breakpoint region from chromosome 8), 2A7 (longest 5′RACE clone), water control. Molecular size standards are indicated in bp. (*B*) The same hybrid and YAC DNAs used in *A* were digested with *Eco*RI and Southern-blotted. The filter was hybridized with a 1.4-kb *TRC8* cDNA fragment that derives from the 3′ end. The normal human *TRC8* fragment is >15 kb, which is reduced to ≈12 kb by the translocation (arrow). The cross-hybridizing fragment in hamster DNA (lanes 2–6) is 8 kb.

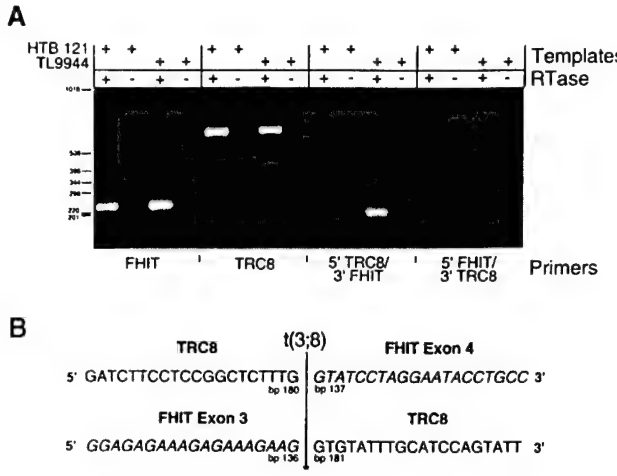


FIG. 4. (*A*) RT-PCR analysis of fusion product expression. RNAs isolated from the t(3;8) lymphoblastoid cell line TL9944 (36) and from a control breast carcinoma cell line (HTB121) were treated with or without RT, as indicated (+ or −) and analyzed for expression of *FHIT* and *TRC8* by PCR. Four primers specific for 5′ and 3′ portions of each gene, F1 and R1 for *FHIT* and R-M and EMR for *TRC8*, were used in combination to detect both wild-type and putative chimeric transcripts. The *FHIT* primer pair generated a product of the expected size (231 bp) as did the *TRC8* primer pair (651 bp). Reciprocal chimeric products were amplified by using R-M plus R1 for 5′ *TRC8*/3′ *FHIT* and F1 plus EMR for 5′ *FHIT*/3′ *TRC8*. Predicted sizes of the chimeric products are 188 and 694 bp, respectively. (*B*) Sequences of 3;8 chimeric transcripts. The RT-PCR-amplified cDNAs in lanes 11 and 15, corresponding to the reciprocal chimeric transcripts, were purified and sequenced on both strands. Forty bp surrounding the boundary between *FHIT* exons 3 and 4 are shown with *FHIT* sequences italicized. The precise position of the fusion on both *TRC8* and *FHIT* transcripts is indicated.

DISCUSSION

The relationship between the 3;8 translocation and kidney cancer has been a long-term project of this laboratory (27, 36–38) and the source of considerable interest among investigators seeking to identify a 3p TSG. Although the 3p14 *FHIT* gene was postulated to be a tumor suppressor (5, 39, 40), this role was questioned (41) based on its biochemical function as a di-adenosine hydrolase (6) combined with the lack of substantial mutations in tumors and the fact that most reported PCR alterations resulted from low-abundance splicing variations (11, 13–15, 17). Moreover, there is little support for the involvement of *FHIT* in renal cancers (16, 42). Similarly, the reintroduction of *FHIT* into tumorigenic cell lines was inconsistent in suppressing tumors, including the fact that a hydrolase "dead" mutant appeared active (39). Otterson *et al.* (43) introduced *FHIT* into six carcinoma cell lines and observed no effects on proliferation, morphology, cell-cycle kinetics, or tumorigenesis. Also, there was no correlation between the presence of truncated RT-PCR products and FHIT protein. We identified a series of 3p14 deletions, many not involving *FHIT* exons, which overlapped FRA3B in various carcinoma

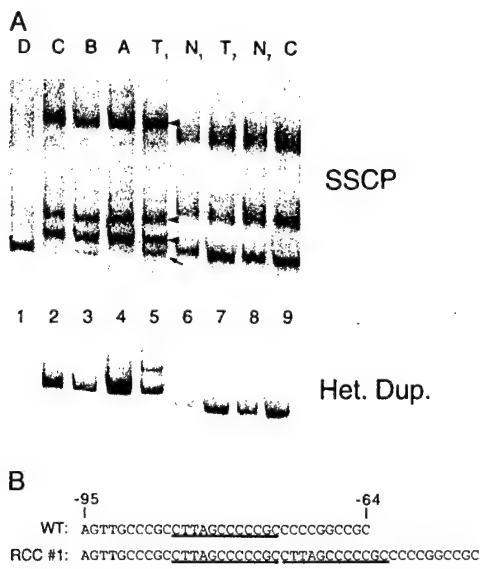


FIG. 5. (*A*) Detection of a tumor-specific somatic mutation by SSCP and heteroduplex analysis. DNA samples were PCR-amplified by using primers flanking the first coding exon of TRC8. The products were denatured, separated on a nondenaturing MDE gel and detected by silver staining. Samples included matched tumor and normal DNAs from patients 1 and 7 (lanes 5–8, respectively) and an unrelated normal control (AG4103, lane 9). A separate SSCP gel was used to isolate four individual SSCP bands from RCC no. 1 (lane 5, marked by an arrow or arrowheads). The excised bands A-D corresponded to the indicated bands in lane 5 from top to bottom. These last templates were reamplified and analyzed by SSCP to determine whether they contained mutant or wild-type sequences. Comparison to lane 5 suggested that bands A and C contained primarily mutant DNA, band B was a mixture of mutant and wild-type, and band D was wild-type only. Results for SSCP (*Upper*) and heteroduplex (*Lower*) are shown. (*B*) RCC no. 1 contains a 12-bp duplication in the 5′ untranslated region. Purified PCR products shown in *A* (lanes 1–4) were sequenced. The mutation consisted of a 12-bp direct duplication (underlined) at bp position −73, which was present in the tumor sample but not in the corresponding normal DNA.

cell lines (14). However, spontaneous deletions also were observed in nontumor backgrounds, and we proposed that the deletions were the result of genomic instability. Although another 3p14 gene might exist, our sequence data totaling 160 kb from FRA3B (14) (plus GenBank updates AF023460 and AF023461) together with 135 kb of nonoverlapping sequence from Inoue *et al.* (44) did not identify any additional definitive genes.

The observation that *FHIT* was involved in a translocation-derived fusion with *HMGIC* (19), the causative gene in a variety of benign tumors, led us to consider that the same event might have occurred in the 3;8 translocation. By using RACE, we were able to identify an additional gene, *TRC8*, from the 8q24 breakpoint. *TRC8* encodes a predicted 664-aa, multitransmembrane protein with similarity to *patched*. This similarity includes the second extracellular domain of *patched*, which is involved in binding sonic hedgehog, as well as its putative SSD. In addition, the first 480 amino acids of *TRC8* and amino acids 440-1100 of *patched* share an organizational similarity (Fig. 1*B*). This similarity begins with the common SSD, followed by the divergent region that is nonconserved among *patched* homologues (45), and finally by the conserved second extracellular loop. *TRC8* lacks the first extracellular loop of *patched* and likewise shows no similarity after the second extracellular loop. Therefore, although *TRC8* has similarity to *patched* and is predicted to be a plasma membrane protein by PSORT (46), it is not the type of direct homologue as is the *Patched 2* gene, for instance (45).

TRC8 may be the critical gene in the 3;8 translocation based on the following: (*i*) its similarity to *patched*, which in turn is responsible for the hereditary basal cell carcinoma syndrome (23, 25), (*ii*) the preservation and expression of *FHIT* coding sequences in 3;8 translocation containing cells (in contrast to the disruption of *TRC8* coding sequences), and (*iii*) its demonstrated mutation in a sporadic renal carcinoma. How *TRC8* may function in human tumorigenesis is presently unknown. In human basal cell carcinomas, *PATCHED* appears to function as a recessive TSG (reviewed in ref. 47). However, medulloblastomas (or perhaps the precursor cell proliferations) occurring in mice heterozygous for a *ptc* mutation might develop without the requirement for loss of the remaining allele (24, 47). In tumors from the 3;8 translocation family, we do not know whether the wild-type allele was lost. A screen of *TRC8* in sporadic renal carcinomas by SSCP demonstrated a mutation in one tumor (Fig. 5*A*). In that sample, very little of the wild-type heteroduplex product can be seen. This rearrangement resulted in an insertion of 12 bp in the tumor, which was not present in the corresponding normal DNA of that patient. This insertion occurs in a consistently predicted stem-loop structure in the 5′ untranslated region. The consequence of this insertion conceivably affects either transcription or translation. Although the frequency of *TRC8* mutations in spontaneous tumors appears low, it is possible this finding is reminiscent of the mutation frequencies observed in BRCA1 and BRCA2 (48, 49). We believe the most exciting aspect of our findings is that *TRC8* appears to define an additional mutation pathway in renal and thyroid cancers. Whether it functions in some analogous way to *patched* or is involved in other developmental processes remains to be determined.

We thank Dr. C. Korch and C. Dessev for their assistance with nucleotide sequencing through the auspices of the University of Colorado Cancer Center DNA Sequencing Core (supported by P30 CA 46934 from the National Institutes of Health) and A. Yokomizo for isolation of tumor DNAs. We also thank Dr. R. S. Brown and the members of the hereditary kidney cancer family. This investigation was supported by Grants CA58187 and DAMD17-94J-4391 from the National Institutes of Health and Department of Defense.

1. Cohen, A. J., Li, F. P., Berg, S., Marchetto, D. J., Tsai, S., Jacobs, S. C. & Brown, R. S. (1979) *N. Engl. J. Med.* **301,** 592–595.
2. Li, F. P., Decker, H. J., Zbar, B., Stanton, V. P., Jr., Kovacs, G., Seizinger, B. R., Aburatani, H., Sandberg, A. A., Berg, S. & Hosoe, S. (1993) *Ann. Intern. Med* **118,** 106–111.
3. Latif, F., Tory, K., Gnarra, J., Yao, M., Duh, F. M., Orcutt, M. L., Stackhouse, T., Kuzmin, I., Modi, W., Geil, L., *et al.* (1993) *Science* **260,** 1317–1320.
4. van den Berg, A. & Buys, C. H. (1997) *Genes Chromosomes Cancer* **19,** 59–76.
5. Ohta, M., Inoue, H., Cotticelli, M. G., Kastury, K., Baffa, R., Palazzo, J., Siprashvili, Z., Mori, M., McCue, P., Druck, T., *et al.* (1996) *Cell* **84,** 587–597.
6. Barnes, L. D., Garrison, P. N., Siprashvili, Z., Guranowski, A., Robinson, A. K., Ingram, S. W., Croce, C. M., Ohta, M. & Huebner, K. (1996) *Biochemistry* **35,** 11529–11535.
7. Sozzi, G., Veronese, M. L., Negrini, M., Baffa, R., Cotticelli, M. G., Inoue, H., Pilotti, S., De Gregorio, L., Pastorino, U., Pierotti, M. A., *et al.* (1996) *Cell* **85,** 17–26.
8. Virgilio, L., Shuster, M., Gollin, S. M., Veronese, M. L., Ohta, M., Huebner, K. & Croce, C. M. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 9770–9775.
9. Negrini, M., Monaco, C., Vorechovsky, I., Ohta, M., Druck, T., Baffa, R., Huebner, K. & Croce, C. M. (1996) *Cancer Res.* **56,** 3173–3179.
10. Sozzi, G., Alder, H., Tornielli, S., Corletto, V., Baffa, R., Veronese, M. L., Negrini, M., Pilotti, S., Pierotti, M. A., Huebner, K. & Croce, C. M. (1996) *Cancer Res.* **56,** 2472–2474.
11. Thiagalingam, S., Lisitsyn, N. A., Hamaguchi, M., Wigler, M. H., Willson, J. K. V., Markowitz, S. D., Leach, F. S., Kinzler, K. W. & Vogelstein, B. (1996) *Cancer Res.* **56,** 2936–2939.

12. Fong, K. M., Biesterveld, E. J., Virmani, A., Wistuba, I., Sekido, Y., Bader, S. A., Ahmadian, M., Ong, S. T., Rassool, F. V., Zimmerman, P. V., *et al.* (1997) *Cancer Res.* **57,** 2256–2267.
13. Gayther, S. A., Barski, P., Batley, S. J., Limin, L., de Foy, K. A. F., Cohen, S. N., Ponder, B. A. J. & Caldas, C. (1997) *Oncogene* **15,** 2119–2126.
14. Boldog, F., Gemmill, R. M., West, J., Robinson, M., Robinson, L., Li, E., Roche, J., Todd, S., Waggoner, B., Lundstrom, R., *et al.* (1997) *Hum. Mol. Genet.* **6,** 193–203.
15. Panagopoulos, I., Thelin, S., Mertens, F., Mitelman, F. & Aman, P. (1997) *Genes Chromosomes Cancer* **19,** 215–219.
16. van den Berg, A., Draaijers, T. G., Kok, K., Timmer, T., van der Veen, A. Y., Veldhuis, P. M., de Leij, L., Gerhartz, C. D., Naylor, S. L., Smith, D. I. & Buys, C. H. (1997) *Genes Chromosomes Cancer* **19,** 220–227.
17. Latil, A., Bièche, I., Fournier, G., Cussenot, O., Pesche, S. & Lidereau, R. (1998) *Oncogene* **16,** 1863–1868.
18. Muller, C. Y., O'Boyle, J. D., Fong, K. M., Wistuba, I. I., Biesterveld, E., Ahmadian, M., Miller, D. S., Gazdar, A. F. & Minna, J. D. (1998) *J. Natl. Cancer Inst.* **90,** 433–439.
19. Geurts, J. M., Schoenmakers, E. F., Roijer, E., Stenman, G. & Van de Ven, W. J. (1997) *Cancer Res.* **57,** 13–17.
20. Frohman, M. A. (1994) *PCR Methods Appl.* **4,** S40–S58.
21. Marigo, V., Davey, R. A., Zuo, Y., Cunningham, J. M. & Tabin, C. J. (1996) *Nature (London)* **384,** 176–179.
22. Stone, D. M., Hynes, M., Armanini, M., Swanson, T. A., Gu, Q., Johnson, R. L., Scott, M. P., Pennica, D., Goddard, A., Phillips, H., *et al.* (1996) *Nature (London)* **384,** 129–134.
23. Hahn, H., Wicking, C., Zaphiropoulous, P. G., Gailani, M. R., Shanley, S., Chidambaram, A. & Vorechovsky, I. (1996) *Cell* **85,** 841–851.
24. Goodrich, L. V., Milenkovic, L., Higgins, K. M. & Scott, M. P. (1997) *Science* **277,** 1109–1113.
25. Johnson, R. L., Rothman, A. L., Xie, J., Goodrich, L. V., Bare, J. W., Bonifas, J. M., Quinn, A. G., Myers, R. M., Cox, D. R., Epstein, E. H., Jr. & Scott, M. P. (1996) *Science* **272,** 1668–1671.
26. Drabkin, H. A., Smith, D., Jones, C., Jonsen, M., Sage, M., Gold, S., Glover, T., Dobrovic, A., Bradley, W. E. & Gemmill, R. (1989) *Cancer Cells* **7,** 63–68.
27. Boldog, F. L., Gemmill, R. M., Wilke, C. M., Glover, T. W., Nilsson, A. S., Chandrasekharappa, S. C., Brown, R. S., Li, F. P. & Drabkin, H. A. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 8509–8513.
28. Frohman, M. A. (1993) *Methods Enzymol.* **218,** 340–356.
29. Spritz, R. A., Holmes, S. A., Ramesar, R., Greenberg, J., Curtis, D. & Beighton, P. (1992) *Am. J. Hum. Genet.* **51,** 1058–1065.
30. Henke, W., Herdel, K., Jung, K., Schnorr, D. & Loening, S. A. (1997) *Nucleic Acids Res.* **25,** 3957–3958.
31. Hua, X., Nohturfft, A., Goldstein, J. L. & Brown, M. S. (1996) *Cell* **87,** 415–426.
32. Loftus, S. K., Morris, J. A., Carstea, E. D., Gu, J. Z., Cummings, C., Brown, A., Ellison, J., Ohno, K., Rosenfeld, M. A., Tagle, D. A., *et al.* (1997) *Science* **277,** 232–235.
33. Carstea, E. D., Morris, J. A., Coleman, K. G., Loftus, S. K., Zhang, D., Cummings, C., Gu, J., Rosenfeld, M. A., Pavan, W. J., Krizman, D. B., *et al.* (1997) *Science* **277,** 228–231.
34. Chin, D. J., Gil, G., Russell, D. W., Liscum, L., Luskey, K. L., Basu, S. K., Okayama, H., Berg, P., Goldstein, J. L. & Brown, M. S. (1984) *Nature (London)* **308,** 613–617.
35. Freemont, P. S. (1993) *Ann. N.Y. Acad. Sci.* **684,** 174–192.
36. Gemmill, R. M., Coyle-Morris, J., Ware-Uribe, L., Pearson, N., Hecht, F., Brown, R. S., Li, F. P. & Drabkin, H. A. (1989) *Genomics* **4,** 28–35.
37. Boldog, F. L., Waggoner, B., Glover, T. W., Chumakov, I., Le Paslier, D., Cohen, D., Gemmill, R. M. & Drabkin, H. A. (1994) *Genes Chromosomes Cancer* **11,** 216–221.
38. Drabkin, H. A., Bradley, C., Hart, I., Bleskan, J., Li, F. P. & Patterson, D. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 6980–6984.
39. Siprashvili, Z., Sozzi, G., Barnes, L. D., McCue, P., Robinson, A. K., Eryomin, V., Sard, L., Tagliabue, E., Greco, A., Fusetti, L., *et al.* (1997) *Proc. Natl. Acad. Sci. USA* **94,** 13771–13776.
40. Huebner, K., Hadaczek, P., Siprashvili, Z., Druck, T. & Croce, C. M. (1997) *Biochim. Biophys. Acta* **1332,** M65–M70.
41. Le Beau, M. M., Drabkin, H., Glover, T. W., Gemmil, R., Rassool, F. V., McKeithan, T. W. & Smith, D. I. (1998) *Genes Chromosomes Cancer* **21,** 281–289.
42. Bugert, P., Wilhelm, M. & Kovacs, G. (1997) *Genes Chromosomes Cancer* **20,** 9–15.
43. Otterson, G. A., Xiao, G. H., Geradts, J., Jin, F., Chen, W. D., Niklinska, W., Kaye, F. J. & Yeung, R. S. (1998) *J. Natl. Cancer Inst.* **90,** 426–432.
44. Inoue, H., Ishii, H., Alder, H., Snyder, E., Druck, T., Huebner, K. & Croce, C. M. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 14584–14589.
45. Motoyama, J., Takabatake, T., Takeshima, K. & Hui, C. (1998) *Nat. Genet.* **18,** 104–106.
46. Nakai, K. & Kanehisa, M. (1992) *Genomics* **14,** 897–911.
47. Ingham, P. W. (1998) *Curr. Opin. Genet. Dev.* **8,** 88–94.
48. Futreal, P. A., Liu, Q., Shattuck-Eidens, D., Cochran, C., Harshman, K., Tavtigian, S., Bennett, L. M., Haugen-Strano, A., Swensen, J. & Miki, Y. (1994) *Science* **266,** 120–122.
49. Lancaster, J. M., Wooster, R., Mangion, J., Phelan, C. M., Cochran, C., Gumbs, C., Seal, S., Barfoot, R., Collins, N., Bignell, G., *et al.* (1996) *Nat. Genet.* **13,** 238–240.

The Journal of Neuroscience, June 1, 1999, 19(11):4428–4436

Chemoattraction and Chemorepulsion of Olfactory Bulb Axons by Different Secreted Semaphorins

Fernando de Castro,[1] Lingjia Hu,[2] Harry Drabkin,[2] Constantino Sotelo,[1] and Alain Chédotal[1]

[1]*Institut National de la Santé et de la Recherche Médicale U106, Hôpital de la Salpêtrière, 75013 Paris, France, and* [2]*University of Colorado Health Sciences Center, Denver, Colorado 80262*

During development, growth cones can be guided at a distance by diffusible factors, which are attractants and/or repellents. The semaphorins are the largest family of repulsive axon guidance molecules. Secreted semaphorins bind neuropilin receptors and repel sensory, sympathetic, motor, and forebrain axons. We found that in rat embryos, the olfactory epithelium releases a diffusible factor that repels olfactory bulb axons. In addition, Sema A and Sema IV, but not Sema III, Sema E, or Sema H, are able to orient *in vitro* the growth of olfactory bulb axons; Sema IV has a strong repulsive action, whereas Sema A appears to attract those axons. The expression patterns of *sema A* and *sema IV* in the developing olfactory system confirm that they may play a cooperative role in the formation of the lateral olfactory tract. This also represents a further evidence for a chemoattractive function of secreted semaphorins.

Key words: semaphorin; olfactory system; chemorepulsion; chemoattraction; development; neuropilin

The organization of axonal projections in the rodent olfactory system has been extensively characterized. Axons from olfactory receptor neurons in the olfactory epithelium project ipsilaterally to glomeruli in the main olfactory bulb, where they synapse on the dendrites of the mitral and tufted cells. These neurons project ipsilaterally to the anterior olfactory nucleus and to higher olfactory centers, including the piriform and entorhinal cortex, and some amygdaloid nuclei, collectively referred to as the primary olfactory cortex (Shipley and Ennis, 1996). The axons of the mitral and tufted cells are located immediately under the pial surface (Derer et al., 1977; Schwob and Price, 1984; Marchand and Bélanger, 1991) and form the lateral olfactory tract (LOT).

The molecular mechanisms governing the establishment of axonal projections from the olfactory receptor neurons to the olfactory bulb have started to be unraveled (Wang et al., 1998). In contrast, the development of bulbofugal projections is still poorly understood, although putative guidepost cells for LOT axons have been described (Sato et al., 1998). In the rat embryo, isolated fibers start to leave the olfactory bulb by embryonic day 14 (E14), and by E15 a LOT has clearly formed (Marchand and Bélanger, 1991; Pini, 1993; López-Mascaraque et al., 1996). At this stage of development, the great majority of postmitotic neurons in the olfactory bulb are mitral cells (Bayer, 1983). Therefore, the early LOT is almost solely composed of mitral cell axons. Organotypic co-cultures of olfactory bulb and telencephalic vesicles or membranes has shown that the telencephalon contains precisely localized, short-range positional cues that guide LOT axons (Sugisaki et al., 1995; Hirata and Fujisawa, 1997). The results of other *in vitro* assays (Pini, 1993; Hu and Rutishauser, 1996) suggest that developing rat LOT axons can also be guided from a distance by some unidentified diffusible chemorepulsive factors produced by the septum or the neocortex.

Chemotropic factors that can attract and/or repel axons have been described in a variety of systems (Tessier-Lavigne and Goodman, 1996). To date, there are two characterized chemoattractants, netrin-1 (Kennedy et al., 1994) and the hepatocyte growth factor/scatter factor (Ebens et al., 1996). Potential chemorepellents have been identified in two gene families, the netrins and the semaphorins (Tessier-Lavigne and Goodman, 1996; Mark et al., 1997). Five secreted semaphorins are known in rodents, and they have been shown to bind receptors of the neuropilin family (Chen et al., 1997; Feiner et al., 1997; He and Tessier-Lavigne, 1997; Kolodkin et al., 1997; Takahashi et al., 1998).

Several pieces of evidence suggest that secreted semaphorins could play a role in the development of bulbofugal projections. Neuropilin-1 is highly expressed on the axons of embryonic mitral cells, and neuropilin-2 mRNA is present in all components of the olfactory circuitry (Kawakami et al., 1995; Chen et al., 1997). Moreover *sema III/collapsin-1* mRNAs are found in the olfactory bulb and in the LOT pathway (Giger et al., 1996; Shepherd et al., 1996; Kobayashi et al., 1997). This has led us to investigate whether secreted semaphorins can influence the growth of LOT axons.

MATERIALS AND METHODS

Animals. Wistar rats (IFFA Credo, Lyon, France) were used for the culture experiments and for *in situ* hybridization studies. The day on which a vaginal plug was detected was considered E0, and the day of birth was considered postnatal day 0. Pregnant rats were anesthetized with chloral hydrate (350 mg/kg, i.p.).

Cloning of mouse semaphorin IV sequences. Mouse *semaphorin IV* sequences were amplified by RT-PCR on a random-primed template from mouse A9 cells using the following primers: 59H8, 5′-TTC AAC TTC CTG CTC AAC; and 39G5, 5′-GAA GAC CAT GCG AAT ATC, which were obtained from the human *H-sema IV* sequence (Roche et al., 1996). PCR conditions included 35 cycles with an annealing temperature of 55°C with *Taq* DNA polymerase and buffer (Promega, Madison, WI).

Received Nov. 23, 1998; revised March 8, 1999; accepted March 11, 1999.

A.C. and C.S. are supported by Institut National de la Santé et de la Recherche Médicale and Grant BIO4-CT960-774 from the European Community (EC). F.d.C. is supported by EC Fellowship ERB-4001-GT-970077. H.D. is supported by Grants CA68383 and CA58187 from the National Institutes of Health. We thank Dr. D. J. Flanagan for the APtag4 vector and Drs. C. Christensen and A. W. Püschel for providing us with mouse *sema H* and mouse *sema A*, respectively.

Correspondence should be addressed to Dr. Alain Chédotal, Institut National de la Santé et de la Recherche Médicale U106, Hôpital de la Salpêtrière, 47 Boulevard de l'Hôpital, 75013 Paris, France.

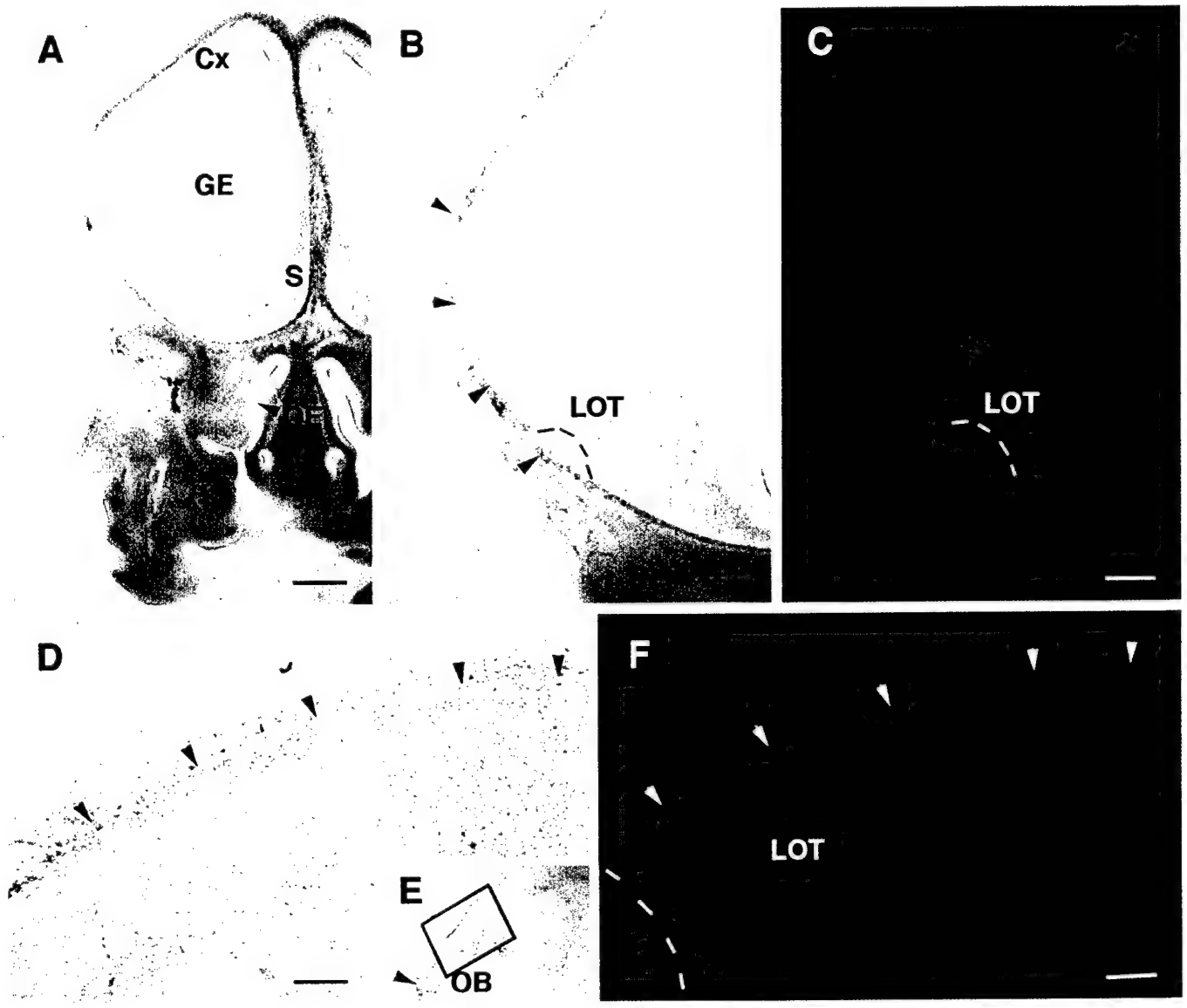


Figure 1. Illustration of the localization of the LOT in rat E15 embryos. Embryos have been injected in the OB with DiI or 4-(4-dihexadecylaminostyryl)-*N*-methylpyridinium iodide (DiA) crystals. *A–C*, Coronal vibratome sections at the level of the septum (*S*), viewed under bright-field (*A, B*) or rhodamine filter. The LOT (*outlined* in *B, C*) grows immediately under the pial surface, in the vicinity of the mesenchyme precursor of the frontal bone (*B, arrowheads*). *D–F*, Horizontal sections corresponding to the area *framed* in *E*. They show, at E15, the rostrocaudal extent of the LOT (labeled axon fascicle in *F*) and the location of the frontal bone (*D, F, arrowheads*). The *dashed line* delimits the caudal extent of the olfactory bulb. The injection site is shown in *E* (*arrowhead*). *Cx*, Cortex; *GE*, ganglionic eminence; *OE*, olfactory epithelium. Scale bars: *A*, 200 μm; *B, C*, 100 μm; *D, F*, 50 μm.

The final products cloned into a T vector prepared from pBluescript II KS and sequenced on an ABI377 through the University of Colorado Cancer Center DNA Sequencing Core. The 0.9-kb product is identical to base pairs 143–1056 of the recently published mouse *H-Sema IV* homolog (Eckhardt and Meyerhans, 1998).

In situ *hybridization.* E14–E15 embryos (five to seven animals each) were perfused with 4% paraformaldehyde in 0.1 M phosphate buffer, pH 7.4 (PFA). Brains were post-fixed in 4% PFA, cryoprotected with 10% sucrose, and sectioned at 20 μm with a cryostat. Antisense riboprobes were labeled with digoxigenin-dUTP (Boehringer Mannheim, Mannheim, Germany) or ^{35}S-UTP (Amersham, Buckinghamshire, UK) as described elsewhere (Chédotal et al., 1998). Controls including hybridization with sense riboprobes prevented hybridization signals.

Explant cultures and co-cultures. The olfactory bulbs from E14–E15 embryos were dissected out as a single piece, and 250–350 μm explants were obtained using fine tungsten needles (Pini, 1993). Only the rostral two-thirds of the bulb were used to try to eliminate most of the accessory olfactory bulb. Explants were co-cultured with aggregates (~750 μm in diameter) of COS cells transfected with secreted alkaline phosphatase (AP), using the AP-Tag-4 vector (a gift from Dr. J. Flanagan, Harvard University, Boston, MA), or human *sema III-myc* (Messersmith et al., 1995), human *sema IV-AP* (Chen et al., 1997), mouse *sema E-AP* (Chen et al., 1997), mouse *sema A-myc*, or mouse *sema H-myc* (Chédotal et al., 1998). COS cell aggregates were prepared using the hanging drop method (Kennedy et al., 1994). Expression was controlled on Western blots using the monoclonal 9E10 anti-myc antibody or a polyclonal anti-AP antibody (Dako, High Wycombe, UK). Explants were embedded in rat tail collagen gel as previously described (Lumsden and Davies, 1986), and cultured for 36–60 hr in DMEM (Seromed, Berlin, Germany) supplemented with L-glutamine, D-glucose, and 10% horse serum (all from Life Technologies, Gaithersburg, MD). Co-cultures were incubated in a 5% CO_2, 37°C, 95% humidity incubator.

Explants were fixed for 1 hr in ice-cold 4% PFA. For the visualization of neuronal processes, cultures were rinsed several times in 0.1 M PBS, blocked with 1% normal goat serum, incubated with a neuron-specific anti-class III β-tubulin monoclonal antibody (1:3000, clone TUJ-1;

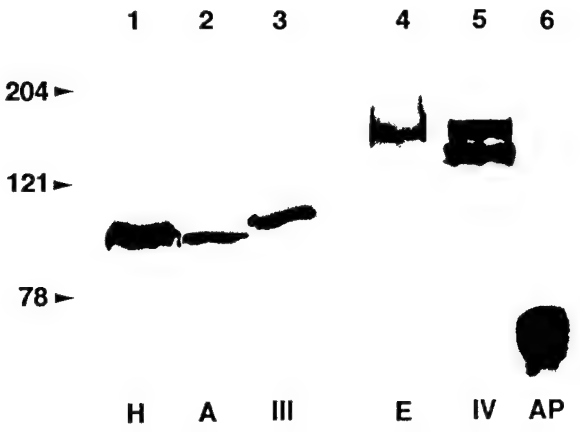


Figure 2. Western blot analysis of epitope-tagged secreted semaphorins. Recombinant proteins were collected in the supernatant of transfected COS cells and analyzed with anti-myc (*lanes 1–3*) or anti-AP (*lanes 4–6*) antibodies. The myc tag is at the C terminus, whereas the AP tag is at the N terminus. In *lane 6*, a single band of ~70 kDa is detected and corresponds to COS cells expressing only secreted AP. For Sema H (*lane 1*), Sema A (*lane 2*), Sema III (*lane 3*), and Sema E-AP (*lane 4*) a single band is observed at their expected molecular weight. In the case of Sema IV-AP (*lane 5*) two bands are detected, the top one at the expected molecular weight of 160 kDa and the bottom one at ~130 kDa, probably corresponding to a partially cleaved protein.

Babco, Richmond, CA; Moody et al., 1989), followed by an HRP-conjugated donkey anti-mouse antibody (1:2000; Jackson Immuno-Research, West Grove, PA), and developed with a diaminobenzidine reaction. Other co-cultures were kept in 4% PFA, and the olfactory bulb explant was injected with a small crystal of lipophilic tracer 1,1′-dioctadecyl-3,3,3′,3′ tetramethylindocarbocyannine (DiI; Molecular Probes, Eugene, OR). After 7–14 d at 37°C in the dark, to allow the diffusion of the tracer, explants were recorded and photographed under rhodamine fluorescent optics.

Quantification. After β-tubulin immunostaining or DiI labeling, the surface covered by the neurites growing from the explant was measured in the proximal and distal quadrants (Wang et al., 1996) using Imstar (Paris, France) software. Microphotographs of each individual explant were digitally scanned with a Nikon CP-9003 camera and transfer to a computer (Imstar). The contour of the area covered by the neuritic processes was acquired by hand with computer-aided filling using a specially devised package from Imstar (see Fig. 4*C*). This measure takes into account both neurite lengths and numbers. In addition, the average length of the neuritic bundles (three to four were used in each quadrant) was measured together with the distance separating the explants from the COS aggregates. Moreover, the mean number of neurites in both the proximal and distal quadrants was also determined (see Table 2). Finally, individual cultures were additionally classified as described in Table 1. Data were statistically analyzed using the Student's *t* test or paired *t* test and their corresponding nonparametric tests (Mann–Whitney and Wilcoxon tests, respectively).

RESULTS

Early development of the rat lateral olfactory tract

At E14–E15 the olfactory bulb can be clearly distinguished at the rostral end of the forebrain, allowing us to inject tiny DiI crystals in the bulb primordium (Fig. 1*E*). Such injections lead to the anterograde tracing of a compact axonal bundle that runs rostrocaudally just under the pial surface (Fig. 1) and corresponds to the presumptive LOT. Dorsoventrally, the LOT is at equal distance from the cortex and the septum (Fig. 1*A–C*). Laterally, the LOT is near a relatively thin leaflet of mesenchymal cells that represents the precursors of the frontal bone (Fig. 1*B,D,F, arrowheads*). They are so closed that the lipophilic dye can, in some cases, diffuse from the LOT to the mesenchyme (Fig. 1*C*). The schematic representation in Figure 7 summarizes the embryonic position of olfactory bulb projecting neurons (in *red*) and possible interactions with closely related cells.

LOT axons are repelled by Sema IV and attracted by Sema A

To test directly whether secreted semaphorins influence the growth of bulbofugal axons, we cultured olfactory bulb explants (from E14–E15 rat embryos) with COS cells that had been transiently transfected with expression constructs for all five known mammalian secreted semaphorins (A, III/D, E, IV, and H). Explants were cultured for 36–60 hr, fixed, and stained with an anti-class III β-tubulin monoclonal antibody that labels the entire population of axons growing from the explant (Moody et al., 1989). The expression of the diverse epitope-tagged semaphorins was verified by Western blotting (Fig. 2). Under nonreducing conditions, all secreted semaphorins, with the exception of Sema IV (see below), run as a single band (at the expected molecular weight of ~95 kDa for myc-tagged Sema A, Sema III, and Sema H, and ~160 kDa for AP-tagged Sema E). In the case of Sema IV two bands were observed, one at ~160 kDa and another one at ~130 kDa, which could represent a partially cleaved fragment. This suggests that in COS cells, most of the secreted semaphorins are not cleaved before secretion, contrary to what has been observed using 293T cells (Adams et al., 1997). Observations similar to ours were made using an Fc-tagged collapsin-1 (Eickholt et al., 1997).

We found that olfactory bulb axons grow symmetrically when confronted with COS cells mock-transfected or transfected with an alkaline phosphatase expression construct (Figs. 3*A*, 4; Table 1). We could not detect any effect of COS cells secreting Sema III, Sema E, or Sema H (Fig. 3*B–D*, Table 1). In all these cases the surface covered by neurites was similar in the proximal and distal quadrants. In contrast, olfactory bulb axons were strongly repelled when confronted with COS cells expressing Sema IV (Figs. 3*E*, 4; Table 1). Repulsion could be clearly detected after 36 hr *in vitro* but was stronger after 60 hr; therefore, quantification was done at that stage. The area covered by the neurites and the mean number of neurites in the distal quadrant were more than two times larger than in the proximal quadrant (Fig. 4*B*, Table 2). Neurites were also much longer in the distal (372.4 ± 15.7 μm; $n = 83$) than in the proximal (234.8 ± 9.8 μm; $n = 83$) quadrant, suggesting that Sema IV somewhat reduces the growth of olfactory bulb (OB) axons. Moreover, a strong or moderate repulsion was observed in 54% of the cases (Fig. 4*A*, Table 1).

Neurite outgrowth was also asymmetrical in the case of explants facing Sema A-expressing COS cell aggregates. Surprisingly, Sema A appeared to exert a moderately attractive effect on olfactory bulb neurites (Figs. 3*F*, *4*; Table 1). The neurite surface in the proximal quadrant was significantly higher of 33% when compared with the distal one (Fig. 4*B*), and the mean number of neurites in the proximal quadrant was very significantly increased (Table 2). In addition, chemoattraction was observed in 40% of the explants in comparison with 11% of control COS cells or 0% of Sema IV-expressing cells (Fig. 4*B*, Table 1). Repulsion with Sema A was observed in merely 9% of the explants (Fig. 4*A*, Table 1). This attractive effect is comparable with the one that has been recently shown for Sema E on neocortical axons (Bagnard et al., 1998). In addition, the average length of neurite bundles was comparable in the proximal (273 ± 8 μm; $n = 92$) and distal (231 ± 10 μm; $n = 92$) quadrants and similar to the average distance (266 ± 10 μm; $n = 92$) between the explants and the COS aggregates. This shows that the Sema A has essentially

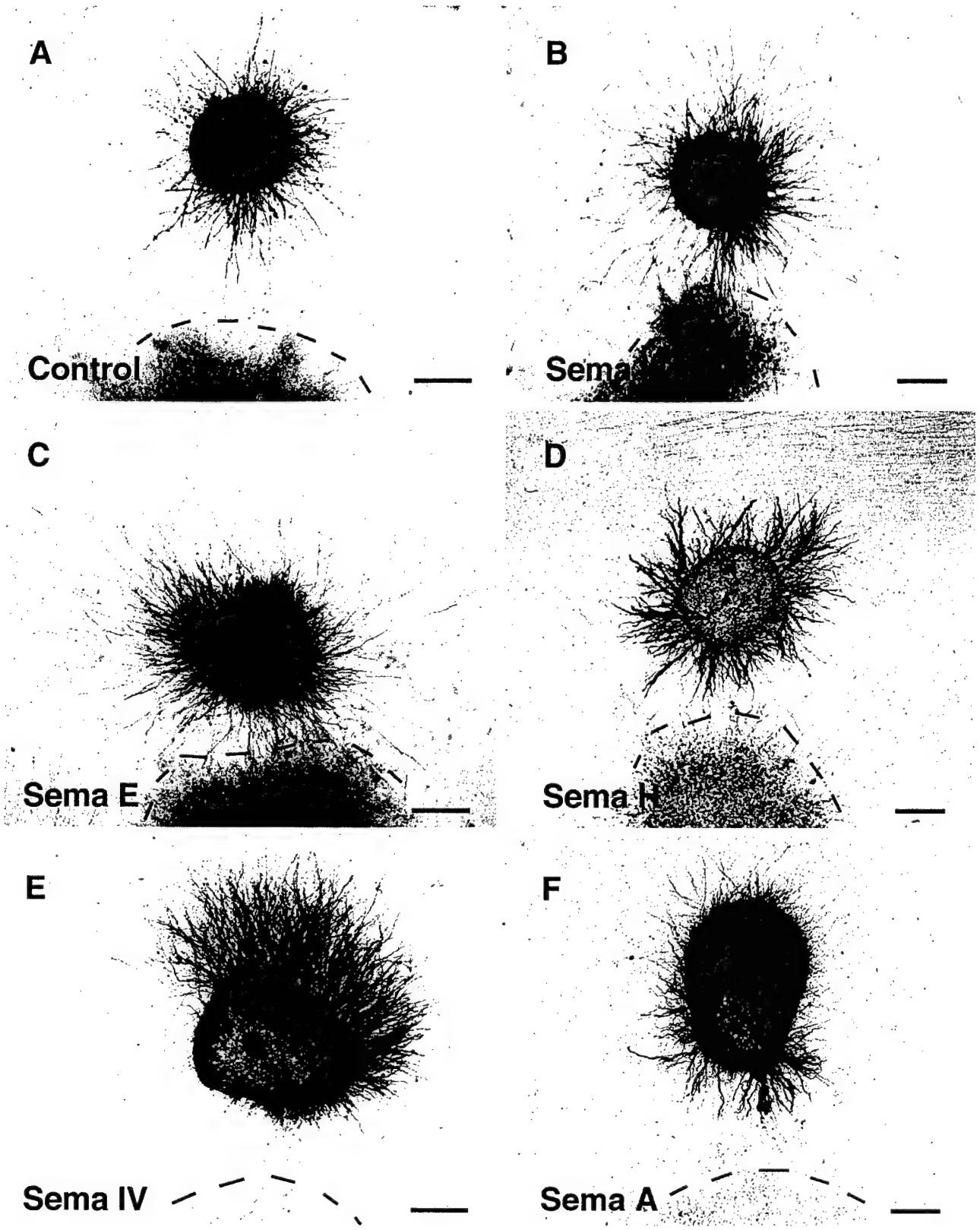


Figure 3. Illustration of the co-culture experiments. E14–E15 rat olfactory bulb explants were co-cultured 60 hr next to aggregates of control COS cells (*A*) or COS cells transfected with *H-sema III* (*B*), mouse *sema E* (*C*), mouse *sema H* (*D*), *H-sema IV* (*E*), or mouse *sema A* (*F*). All explants were fixed and stained with anti-β-tubulin antibodies. Olfactory bulb axons grow symmetrically, in the case of control cells (*A*) and COS cells expressing Sema III (*B*), Sema E (*C*), and Sema H (*D*), whereas they are strongly repelled by Sema IV-expressing COS cell aggregates (*E*). With Sema A-expressing cells (*F*), olfactory bulb axons grow in all quadrants, but they are more numerous in the quadrant facing the COS cells aggregate, indicative of an attraction. Scale bars: *A–D*, 170 μm; *E, F*, 150 μm.

Table 1. Semiquantitative evaluation of the effect of secreted Semaphorins and olfactory epithelium (OE) on axonal outgrowth of olfactory bulb explants

		Axonal outgrowth				
Coculture condition	*n*	++	+	=	−	−−
COS control	71	3	5	51	8	4
Sema III	34		2	28	4	
Sema E	48		5	36	7	
Sema H	24		3	18	3	
Sema A	109	15	27	57	5	5
Sema IV	96			44	16	36
OE	56	1	9	17	11	18

Explants were labeled using anti-β-tubulin antibodies. Co-cultures were classified as follows: ++, strong axonal attraction (surface occupied by axonal growth is more than threefold larger in the proximal than in the distal quadrant); +, moderate attraction of axons (axonal surface is twofold to threefold larger in the proximal than in the distal quadrant); =, radial axonal growth (axonal surface in the proximal and distal quadrants differed by less than twofold); −, moderate axonal repulsion (neurite surface twofold to threefold larger in the distal than in the proximal quadrant); −−, strong axonal repulsion (surface occupied by axons is more than threefold broader in the distal than in the proximal quadrant).

a directional effect and does not stimulate the growth. Repulsive secreted semaphorins also have mostly a directional action (Messersmith et al., 1995; Song et al., 1998).

Sema A and Sema IV expression in the rat olfactory system

We tried to determine the possible function of Sema A and Sema IV in the development of the LOT *in vivo* by studying the expression pattern of their mRNAs in E14–E15 rat embryos (Fig. 5). At E15, *sema A* mRNA was not detected in the forebrain (Fig. 5*A,B*), as previously mentioned (Püschel et al., 1996). Nevertheless, we could observe a strong expression of *sema A* in the mesenchyme precursor of the frontal bone extending from the base of the olfactory bulb rostrally to the piriform cortex caudally (Fig. 5*A,B*). Therefore, *sema A* expression is restricted to cells that are juxtaposed to the pathway followed by LOT axons. *sema IV* was expressed in the cortical plate of the neocortex, as previously described (Chédotal et al., 1998), and in the ganglionic eminence (Fig. 5*C,D*). High *sema IV* mRNA expression was also observed in the olfactory epithelium (Fig. 5*E*) and was absent from the septum (Fig. 5*C,D*). In addition, at E15, *sema IV* mRNA was not expressed in the olfactory bulb (results not shown).

Chemorepulsion of olfactory bulb axons by the olfactory epithelium

Because Sema IV repels LOT axons and is highly expressed in the olfactory epithelium, we tried to determine whether we could show long-range effects of the olfactory epithelium on olfactory bulb axons. We therefore co-cultured in collagen gel E14–E15 olfactory bulb explants next to olfactory epithelial explants. In 52% of the cases, the growth of olfactory bulb axons, visualized with DiI, was opposed to the epithelial explant (Figs. 4*A*, 6; Table 1), and neurites were significantly more numerous in the distal quadrant than in the proximal one (Table 2). This demonstrates that the embryonic olfactory epithelium secretes a factor repulsive for olfactory bulb axons.

We also tried to co-culture olfactory bulb explants next to explants of mesenchyme surrounding the LOT but could not observe clear attractive or repulsive effects, because in 24 hr mesenchymal cells migrate in the collagen gel and reach and partially cover OB explants (results not shown).

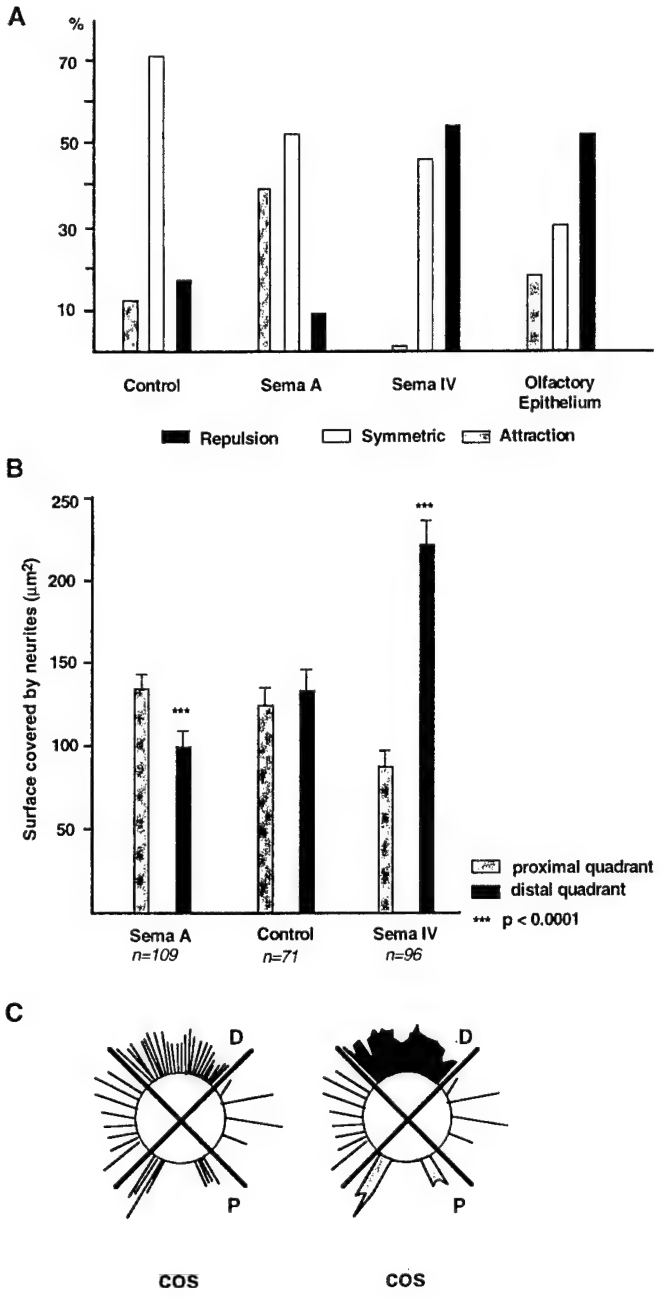


Figure 4. *A*, Histogram summarizing the data shown in Table 1 and illustrating the proportion (percent) of olfactory bulb explants in which axons look attracted, repelled, or unaffected when co-cultured next to control COS cells, Sema A- or Sema IV-expressing cells, and olfactory epithelium. *B*, Histogram showing the axonal surface (mean ± SEM) in the different combinations of olfactory bulb explants and COS cells. All explants were fixed and stained with anti-β-tubulin antibodies. The axonal surface was measured in the distal (*black bars*) and proximal (*gray bars*) quadrants, as indicated in *C*. Olfactory bulb axons are strongly repelled by Sema IV but attracted by Sema A. Control COS cells have no effect. The *asterisks* denote differences between both quadrants (proximal and distal) that are significant at the $p = 0.0001$ level (Mann–Whitney rank sum test and Wilcoxon signed rank test). *C*, Schema of the method used to quantify neurite outgrowth. An image analysis system was used to determine the overall area covered by neurites in the proximal (*P, gray area*) and distal (*D, black area*), as represented on the *right*. This gives an estimate of neurite lengths and numbers.

Table 2. Mean number of neurite bundles in OB explants in different coculture conditions

Coculture conditions	*n*	Mean no. of neurite bundles ± SEM		*p*
		Proximal quadrant	Distal quadrant	
OB + COS control	10	46.3 ± 10	46.7 ± 11	NS (>0.05)
OB + Sema A	10	65.2 ± 18	24.8 ± 7	<0.0001
OB + Sema IV	10	24.9 ± 7	73 ± 11	<0.0001
OB + OE	10	17.9 ± 6	57.8 ± 13	<0.0001

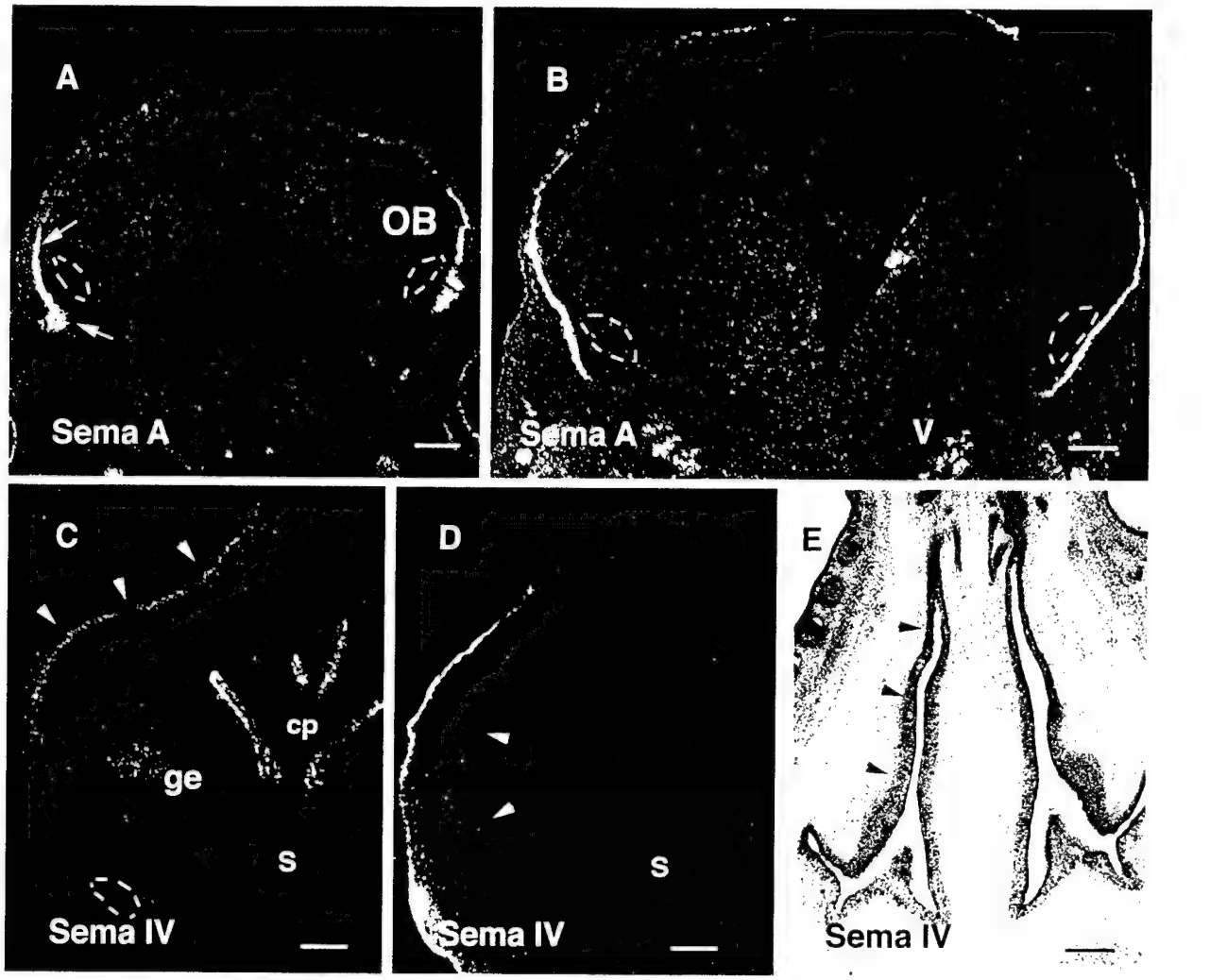


Figure 5. Expression pattern of *sema A* and *sema IV* in the developing head. Hybridizations were performed with ^{35}S-labeled (*A–D*) or digoxigenin-labeled (*E*) riboprobes on coronal (*A–D*) and horizontal (*E*) sections of E15 rat brains. The prospective location of the LOT is delimited by *dashes* in *A–C*. *A*, *B*, *sema A* mRNA is found in the mesenchyme precursor of the frontal bone (*A*, *arrows*), immediately adjacent to the LOT. Expression starts at the base of the OB in *A* and runs caudally to the level of the trigeminal ganglion (*V*) that is also expressing *sema A* (see *B*). *C*, *D*, *sema IV* is expressed in the cortical plate (*arrowheads*) and in the ganglionic eminence (*ge*). Strong expression is also found in the choroid plexuses (*cp*) and the skin. Note in *C* and *D* that the septum (*S*) is not expressing *sema IV*. In *E*, *sema IV* expression is homogeneously high in the olfactory epithelium (*arrowheads*). Scale bars: *A*, 215 μm; *B*, *D*, 110 μm; *C*, 240 μm; *E*, and 300 μm.

DISCUSSION

We found that olfactory bulb axons, which will form the LOT, respond to Sema IV and Sema A but not to other secreted semaphorins. In collagen gel assays, Sema IV repels olfactory bulb axons, whereas, more surprisingly, Sema A acts as an attractant. In addition, in co-culture experiments, the olfactory epithelium exercises a repulsive action on the axons growing from olfactory bulb explants. Finally, the expression patterns of *sema A* and *sema IV* in the embryonic olfactory system and in structures next to the lateral olfactory tract strongly suggest that these semaphorins might play a role in the formation of the LOT.

Developing axons are known to be guided from a distance by diffusible molecules that can be attractive, repulsive, or both (Tessier-Lavigne and Goodman, 1996). Our results suggest that

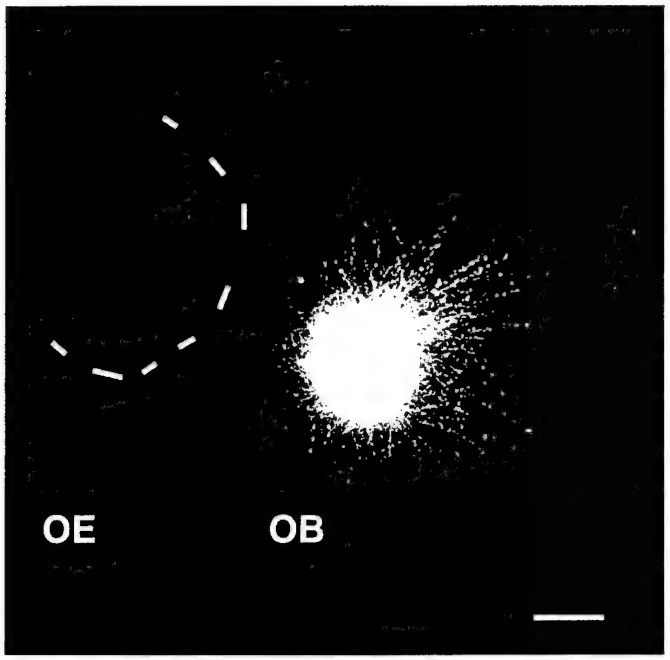


Figure 6. Olfactory bulb and olfactory epithelium cocultures. OB explants were labeled with DiI and placed near an unlabeled explant of olfactory epithelium (*OE, dotted line*). Labeled axons grow almost exclusively on the distal side of the explant opposite to the olfactory epithelium, indicating that the OE releases a diffusible repulsive factor. Scale bar, 200 μm.

chemoattraction and chemorepulsion exert a cooperative action during the development of the LOT. First, we showed that the olfactory epithelium releases a diffusible factor that repels olfactory bulb axons and could force them to leave the olfactory bulb primordium and grow caudally. This repellent effect is mimicked by Sema IV and not by any other secreted semaphorins or by netrin-1, another stong chemorepellent (F. de Castro, A. Chédotal, and C. Sotelo, unpublished observations). The finding that *sema IV* is expressed in the olfactory epithelium (also see Giger et al., 1998) at the time LOT axons grow suggests that this molecule could be the epithelial-derived repellent factor (Fig. 7; see below). In addition, it is known that developing LOT axons never enter the embryonic neocortex (Schwob and Price, 1984), which produces a repulsive factor for LOT axons (Pini, 1993). *sema IV*, whose mRNA is highly expressed in the cortical plate of the neocortex (Chédotal et al., 1998; Giger et al., 1998 and present results), becomes an excellent candidate to be the secreted factor preventing the invasion of the neocortex by LOT axons (Fig. 7).

Chemorepulsion has been observed in several neuronal systems (Fitzgerald et al., 1993; Colomarino and Tessier-Lavigne, 1995; Tamada et al., 1995; Varela-Echavarría et al., 1997; Chédotal et al., 1998; Rochlin and Farbman, 1998; Tuttle et al., 1998), but the first evidence came from the embryonic rat olfactory system in which the septum produces a diffusible factor repelling olfactory bulb axons (Pini, 1993; Fig. 7). This still uncharacterized activity was supposed to explain why LOT axons grow away from the midline (Pini, 1993), although recent *in vitro* results have shown that mitral cell axons also elongate along their normal pathway without a septum (Sugisaki et al., 1995). Because Sema IV and the other known secreted semaphorins have not been found in the embryonic septum, the repulsive factor, produced by the latter, might not belong to the semaphorin family.

In addition, olfactory bulb axons grow preferentially toward aggregates of COS cells producing Sema A but not any other secreted semaphorins, suggesting that chemoattraction could also be involved in the formation of the LOT. In the embryonic rat head, *sema A* mRNA is almost exclusively found in the frontal bone primordium that is apposed to the forebrain parenchyma and therefore immediately adjacent to the LOT, at a distance that could be smaller than 20 μm (Meller and Tetzlaff, 1975). Although direct evidence is still lacking, our results suggest that Sema A released by the mesenchymal cells could attract in the CNS axons of the lateral olfactory tract. At this early stage, the immature meningeal covers are permeable, and the glia limitans is covered by a basal membrane, which prevents axons to leave the CNS. This could explain why LOT axons grow preferentially in the most superficial zone of the forebrain (Fig. 7). Sema A-mediated attraction was weaker than Sema IV repulsion, a difference that could emerge from the heterogeneity of olfactory bulb axons together with a differential action on a large axonal population for Sema IV and on a smaller one for Sema A (similarly, at E14 in the rat, Sema III repels only NGF-sensitive dorsal root ganglion axons but not NT3-sensitive ones; Messersmith et al., 1995). This assumption is based on the fact that mitral and tufted cells are morphologically, neurochemically, and functionally heterogeneous (Macrides et al., 1985; Shipley and Ennis, 1996).

It has recently been shown that Sema E can attract cortical axons (Bagnard et al., 1998). The present results confirm that in the vertebrate CNS, secreted semaphorins could promote the growth of certain axonal populations in normal developmental conditions. In addition, in cultured embryonic *Xenopus* spinal neurons, an experimental elevation of the cytosolic concentration of cGMP can turn Sema III repulsive action into an attraction (Song et al., 1998). Moreover, G-Sema I, a grasshopper transmembrane semaphorin that was previously shown to inhibit the growth of other sensory axons (Kolodkin et al., 1992), appears to promote the growth of axons in the subgenual organ (Wong et al., 1997). Similarly, Sema A, attractive in this study, is repellent for sympathetic axons (Adams et al., 1997; Takahashi et al., 1998). Therefore, we have enough evidence to start thinking that semaphorins might be bifunctional guiding cues, attractive and repulsive, like the midline-derived netrin-1 (Serafini et al., 1994; Colomarino and Tessier-Lavigne, 1995; Varela-Echavarría et al., 1997). Non-neuronal tissues, such as the somites, are known to influence the patterning of axonal projections in the peripheral nervous system (Tannahill et al., 1998). In addition, diffusible factors produced by the the notochord and other mesodermic tissues play a role in neuronal differentiation in the CNS (LaMantia et al., 1993; Rubenstein and Beachy, 1998). To our knowledge, the present results represent the first evidence for a presumptive influence of non-neuronal peripheral tissue on axon tract formation and orientation in the vertebrate CNS.

The identity of the receptors involved in Sema A and Sema IV signaling in olfactory bulb axons is still unknown. The transmembrane proteins neuropilin-1 and -2 have been shown to be receptors, or components of receptors, for vertebrate-secreted semaphorins (Chen et al., 1997; Feiner et al., 1997, He and Tessier-Lavigne, 1997; Kolodkin et al., 1997, Giger et al., 1998; Takahashi et al., 1998). Recently, it has been shown that Sema IV-induced repulsion of sensory axons is mediated by neuropilin-2 (Giger et al., 1998). Moreover, Sema A collapsing activity also requires neuropilin-2, perhaps in combination with neuropilin-1 (Takahashi et al., 1998). Both neuropilins are expressed by mitral neurons and mitral cell axons at the time they leave the olfactory

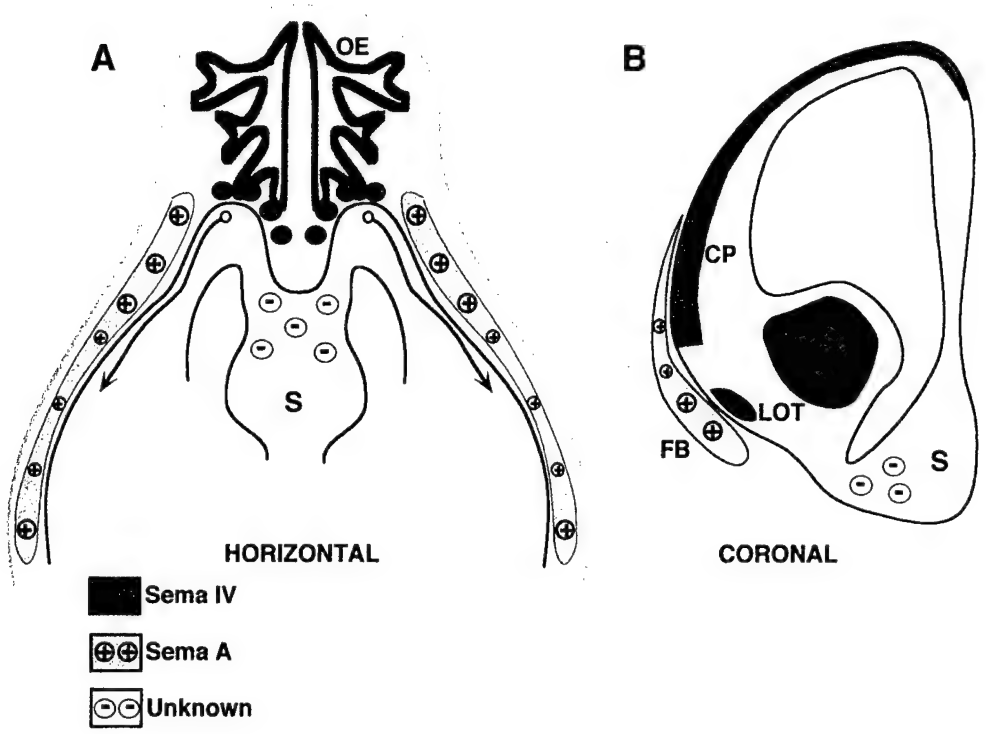


Figure 7. Schematic representation in horizontal (*A*) and coronal (*B*) planes of the chemorepulsive and chemoattractive factors that influence the formation of the lateral olfactory tract. The olfactory epithelium (*OE*) would released a repulsive factor (*violet circles*), probably Sema IV, forcing LOT axons (*red*) to leave the olfactory bulb. An unidentified factor (*green circles*) produced by the septum (*S*) would prevent LOT axons to cross the midline. In addition, LOT axons would grow very superficially, being attracted by Sema A (*orange circles*) secreted by the mesenchyme precursor of the frontal bone (*gray*).

bulb (Sugisaki et al., 1995; Kawakami et al., 1996; Chen et al., 1997; F. de Castro and A. Chédotal, unpublished observations), but in *neuropilin-1* knock-out mice (Kitsukawa et al., 1997) the LOT develops normally, suggesting that neuropilin-2 could be the receptor mediating Sema IV and Sema A responses. It would be interesting to determine whether the repulsive and attractive effects of Sema A are mediated by different receptors. In *Xenopus* axons, it is clear that the same second messengers may be used for both Sema III-mediated attraction and repulsion and that neuropilin-1 is the receptor for both (Song et al., 1998). Finally, the receptor might not be a neuropilin, because it has recently been shown that virus-encoded semaphorin protein receptor, a member of the plexin family of transmembrane proteins, is a receptor for a viral semaphorin (Comeau et al., 1998), and that in *Drosophila*, plexin A binds transmembrane semaphorins (Winberg et al., 1998). Whether plexins are receptors for other vertebrate-secreted or transmembrane semaphorins remains unknown (also see Fujisawa and Kitsukawa, 1998).

Semaphorins have been involved in the development and regeneration of other axonal connections of the olfactory system (Shepherd et al., 1996; Pasterkamp et al., 1998). Collapsin-1/Sema III can induce the collapse of olfactory axonal growth cones (Kobayashi et al., 1997) and could serve as a stop signal for these axons, initially preventing them from invading the olfactory bulb (Giger et al., 1996). Recently it has been shown that Sema A and Sema E can antagonize Sema III, preventing its binding to neuropilin-1 (Takahashi et al., 1998). Olfactory receptor neurons express Sema IV, which also binds neuropilin-1 but does not induce DRG growth cone collapse (A. Chédotal, unpublished observations) and could also be an antagonist of neuropilin-1. In agreement with Takahashi et al. (1998), Sema IV could act as a repellent for olfactory bulb axons and an antirepellent for olfactory axons, allowing them to enter the bulb even in the presence of a high concentration of Sema III in their target territory. Finally, several studies have shown that the formation of the olfactory bulb depends on the olfactory receptor neurons, independently of synapse formation (Graziadei et al., 1978; Stout and Graziadei, 1980; Gong and Shipley, 1995). The factors responsible for these actions have not been identified but it would be interesting to determine whether it could be a semaphorin.

In conclusion, during axonogenesis, LOT axons are probably influenced by multiple diffusible factors, repulsive and attractive, in combination with a variety of short-range cues (Sugisaki et al., 1995; Sato et al., 1998).

REFERENCES

Adams RH, Lohrum M, Klostermann A, Betz H, Püschel AW (1997) The chemorepulsive activity of secreted semaphorins is regulated by furin-dependent proteolytic processing. EMBO J 16:6077–6086.

Bagnard D, Lohrum M, Uziel D, Püschel AW, Bolz J (1998) Semaphorins act as attractive and repulsive guidance signals during the development of cortical projections. Development 125:5043–5053.

Bayer SA (1983) ^{3}H-Thymidine-radiographic studies of neurogenesis in the rat olfactory bulb. Exp Brain Res 50:329–340.

Chédotal A, Del Río JA, Ruiz M, He Z, Borrell V, de Castro F, Ezan F, Goodman CS, Tessier-Lavigne M, Sotelo C, Soriano E (1998) Semaphorins III and IV repel hippocampal axons via two distinct receptors. Development 125:4313–4323.

Chen H, Chédotal A, He Z, Goodman CS, Tessier-Lavigne M (1997) Neuropilin-2, a novel member of the neuropilin family, is a high affinity receptor for the semaphorins sema E and sema IV but not sema III. Neuron 19:547–559.

Colomarino S, Tessier-Lavigne M (1995) The axonal chemoattractant netrin-1 is also a chemorepellent for trochlear motor axons. Cell 81:621–629.

Comeau MR, Johnson R, DuBose RF, Petersen M, Gearing P, VandenBos T, Park L, Farrah T, Buller RM, Cohen JI, Stockbine LD, Rauch C, Spriggs MK (1998) A poxvirus-encoded semaphorin induces cytokine production from monocytes and binds to a novel cellular semaphorin receptor, VESPR. Immunity 8:473–482.

Derer P, Caviness VS, Sidman RL (1977) Early cortical histogenesis in the primary olfactory cortex of the mouse. Brain Res 123:27–40.

Ebens A, Brose K, Leonardo ED, Hanson Jr MG, Bladt F, Birchmeier C, Barres BA, Tessier-Lavigne M (1996) Hepatocyte growth factor/scat-

ter factor is an axonal chemoattractant and a neurotrophic factor for spinal motorneurons. Neuron 17:1157–1172.
Eckhardt F, Meyerhans A (1998) Cloning and expression pattern of a murine semaphorin homologous to H-sema IV. NeuroReport 9:3975–3979.
Eickholt BJ, Morrow R, Walsh FS, Doherty P (1997) Strutural features of collapsin required for biological activity and distribution of binding sites in the developing chick. Mol Cell Neurosci 9:358–371.
Feiner L, Koppel AM, Kobayashi H, Raper JA (1997) Secreted chick semaphorins bind recombinant neuropilin with similar affinities but bind different subsets of neurons *in situ*. Neuron 19:539–545.
Fitzgerald M, Kwiat GC, Middleton J, Pini A (1993) Ventral spinal cord inhibition of neurite outgrowth from embryonic rat dorsal root ganglia. Development 117:1377–1384.
Fujisawa H, Kitsukawa T (1998) Receptors for collapsin/semaphorins. Curr Opin Neurobiol 8:587–592.
Giger RJ, Wolfer DP, De Wit GMJ, Verhaagen J (1996) Anatomy of rat semaphorin III/collapsin-1 mRNA expression and relationship to developing nerve tracts during neuroembryogenesis. J Comp Neurol 375:378–392.
Giger RJ, Urquhart ER, Gillespie SKH, Levengood DV, Ginty DD, Kolodkin AL (1998) Neuropilin-2 is a receptor for semaphorin IV: insight into the structural basis of receptor function and specificity. Neuron 21:1079–1092.
Gong Q, Shipley MT (1995) Evidence that pioneer olfactory axons regulate telencephalon cell kinetics to induce the formation of the olfactory bulb. Neuron 14:91–101.
Graziadei PPC, Levine RR, Monti-Graziadei GA (1978) Regeneration of olfactory axons and synapse formation in the forebrain after bulbectomy in neonatal mice. Proc Natl Acad Sci USA 75:5230–5234.
He Z, Tessier-Lavigne M (1997) Neuropilin is a receptor for the axonal chemorepellent semaphorin III. Cell 90:739–751.
Hirata T, Fujisawa H (1997) Cortex-specific distribution of membrane-bound factors that promote neurite outgrowth of mitral cells in culture. J Neurobiol 32:415–425.
Hu H, Rutishauser U (1996) A septum-derived chemorepulsive factor for migrating olfactory interneuron precursors. Neuron 16:933–940.
Kawakami A, Kitsukawa T, Takagi S, Fujisawa H (1995) Developmentally regulated expression of a cell surface protein, neuropilin, in the mouse nervous system. J Neurobiol 29:1–17.
Kennedy TE, Serafini T, de la Torre JR, Tessier-Lavigne M (1994) Netrins are diffusible chemotropic factors for commisural axons in the embryonic spinal cord. Cell 78:425–435.
Kitsukawa T, Shimizu M, Sanbo M, Hirata T, Taniguchi M, Bekku Y, Yagi T, Fujisawa H (1997) Neuropilin-semaphorin III/D-mediated chemorepulsive signals play a crucial role in peripheral nerve projection in mice. Neuron 19:995–1005.
Kobayashi H, Koppel AM, Luo Y, Raper JA (1997) A role for collapsin-1 in olfactory and cranial sensory axon guidance. J Neurosci 17:8339–8352.
Kolodkin AL, Mathes DJ, O'Connor TP, Patel NH, Admon A, Bentley D, Goodman CS (1992) Fascilin IV: sequence, expression, and function during growth cone guidance in the grasshopper embryo. Neuron 9:831–845.
Kolodkin AL, Levengood DV, Rowe EG, Tai YT, Giger RJ, Ginty DD (1997) Neuropilin is a semaphorin III receptor. Cell 90:753–762.
LaMantia AS, Colbert MC, Linney E (1993) Retinoic acid induction and regional differentiation prefigure olfactory pathway formation in the mammalian forebrain. Neuron 10:1035–1048.
López-Mascaraque L, De Carlos JA, Valverde F (1996) Early onset of the rat olfactory bulb projections. Neuroscience 70:255–266.
Lumsden A, Davies AM (1986) Chemotropic effect of specific target epithelium in the development of the mammalian nervous system. Nature 323:538–539.
Macrides F, Schoenfeld TA, Marchand JE, Clancy AN (1985) Evidence for morphologically, neurochemically and functionally heterogeneous classes of mitral and tufted cells in the olfactory bulb. Chem Senses 10:175–202.
Marchand R, Bélanger MT (1991) Ontogenesis of the axonal circuitry associated with the olfactory system in the rat embryo. Neurosci Lett 129:285–290.
Mark MD, Lohrum M, Püschel A (1997) Patterning neuronal connections by chemorepulsion: the semaphorins. Cell Tissue Res 290:299–306.
Meller K, Tetzlaff W (1975) Neuronal migration during the early development of the cerebral cortex. Cell Tissue Res 163:313–325.
Messersmith EK, Leonardo ED, Shatz CJ, Tessier-Lavigne M, Goodman CS, Kolodkin AL (1995) Semaphorin III can function as a selective chemorepellent to pattern sensory projections in the spinal cord. Neuron 14:949–959.
Moody SA, Quigg MS, Frankfurter A (1989) Development of the peripheral trigeminal system in the chick revealed by an isotype-specific anti-β-tubulin monoclonal antibody. J Comp Neurol 346:97–118.
Pasterkamp RJ, De Winter F, Holtmaat AJGD, Verhaagen J (1998) Evidence for a role of the chemorepellent semaphorin III and its receptor neuropilin-1 in the regeneration of primary olfactory axons. J Neurosci 18:9962–9976.
Pini A (1993) Chemorepulsion of axons in the developing mammalian central nervous system. Science 261:95–98.
Püschel AW, Adams RH, Betz H (1996) The sensory innervation of the mouse spinal cord may be patterned by differential expression of and by differential responsiveness to semaphorins. Mol Cell Neurosci 7:419–431.
Roche J, Boldog F, Robinson M, Robinson L, Varella-Garcia M, Swanton M, Waggoner B, Fishel R, Franklin W, Gemmill R, Drabkin H (1996) Distinct 3p21.3 deletions in lung cancer, analysis of deleted genes and identification of a new human semaphorin. Oncogene 12:1289–1297.
Rochlin MW, Farbman AI (1998) Trigeminal ganglion axons are repelled by their presumptive targets. J Neurosci 18:6840–6852.
Rubenstein JLR, Beachy PA (1998) Patterning of the embryonic forebrain. Curr Opin Neurobiol 8:18–26.
Sato Y, Hirata T, Ogawa M, Fujisawa H (1998) Requirement for early generated neurons recognized by monoclonal antibody Lot1 in the formation of lateral olfactory tract. J Neurosci 18:7800–7810.
Schwob JE, Price JL (1984) The development of axonal connections in the central olfactory system of rats. J Comp Neurol 223:177–202.
Serafini T, Kennedy TE, Galko MJ, Mirzayan C, Jessell TM, Tessier-Lavigne M (1994) The netrins define a family of axon outgrowth-promoting proteins homologous to *C. elegans* UNC-6. Cell 78:409–424.
Shepherd I, Luo Y, Raper JA, Chang S (1996) The distribution of collapsin mRNA in the developing chick nervous system. Dev Biol 173:185–199.
Shipley MT, Ennis M (1996) Functional organization of olfactory system. J Neurobiol 30:123–176.
Song H, Ming G, He Z, Lehmann M, McKerracher L, Tessier-Lavigne M, Poo M (1998) Conversion of neuronal growth cone responses from repulsion to attraction by cyclic nucleotides. Science 281:1515–1518.
Stout RP, Graziadei PPC (1980) Influence of the olfactory placode on the development of the brain in *Xenopus laevis* (Daudin)-I. Axonal growth and connections of the transplanted olfactory placode. Neuroscience 5:2175–2186.
Sugisaki N, Hirata T, Naruse Y, Kawakami A, Kitsukawa T, Fujisawa H (1995) Positional cues that are strictly localized in the telencephalon induce preferential growth of mitral cell axons. J Neurobiol 29:127–137.
Takahashi T, Nakamura F, Jin Z, Kalb RG, Strittmatter SM (1998) Semaphorins A and E act as antagonists of neuropilin-1 and agonists of neuropilin-2 receptors. Nat Neurosci 1:487–493.
Tamada A, Shirasaki, R, Murakami F (1995) Floor plate chemoattracts crossed axons and chemorepels uncrossed axons in the vertebrate brain. Neuron 14:1083–1093.
Tannahill D, Cook GMW, Keynes RJ (1998) Axon guidance and somites. Cell Tissue Res 290:275–283.
Tessier-Lavigne M, Goodman CS (1996) The molecular biology of axon guidance. Science 274:1123–1133.
Tuttle R, Braisted JE, Richards LJ, O'Leary DDM (1998) Retinal axon guidance by region-specific cues in diencephalon. Development 125:791–801.
Varela-Echavarría A, Tucker A, Püschel AW, Guthrie S (1997) Motor axon subpopulations are respond differentially to the chemorepellents netrin-1 and semaphorin D. Neuron 18:193–207.
Wang F, Nemes A, Mendelsohn M, Axel R (1998) Odorant receptors govern the formation of a precise topographic map. Cell 93:47–60.
Wang L-C, Rachel RA, Marcus RC, Mason CA (1996) Chemosupression of retinal axon growth by the mouse optic chiasm. Neuron 17:849–862.
Winberg ML, Noordermeer JN, Tamagnone L, Comoglio PM, Spriggs MK, Tessier-Lavigne M, Goodman CS (1998) Plexin A is a neuronal semaphorin receptor that controls axon guidance. Cell 95:903–916.
Wong JTW, Yu WTC, O'Connor TP (1997) Transmembrane grasshopper semaphorin I promotes axon outgrowth in vivo. Development 124:3597–3607.

Oncogene (1999) 18, 2589–2597

http://www.stockton-press.co.uk/onc

DEF-3(g16/NY-LU-12), an RNA binding protein from the 3p21.3 homozygous deletion region in SCLC

Harry A Drabkin*[,1], James D West[1], Marc Hotfilder[2], Yee M Heng[3], Paul Erickson[1], Roser Calvo[1,4], Josep Dalmau[5], Robert M Gemmill[1] and Fred Sablitzky[3]

[1]*Division of Medical Oncology, University of Colorado Health Sciences Center, 4200 E. 9th Ave., Denver, Colorado 80262, USA;* [2]*University Hospital Münster, Paediatric Haematology/Oncology, Department of Experimental Neurooncology, Münster, Germany;* [3]*Department of Medicine, The Windeyer Institute of Medical Sciences, University College London, London, UK;* [4]*Department of Medicine, Hospital de Badalona Germans Trias i Pujol, Barcelona, Spain;* [5]*Laboratory of Molecular Neuro-Oncology, Department of Neurology, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021, USA*

DEF-3(g16/NY-LU-12) encodes a novel RNA binding protein isolated by positional cloning from an SCLC homozygous deletion region in 3p21.3 and, in parallel, as a differentially expressed gene during myelopoiesis from FDCPmix-A4 cells. DEF-3(g16/NY-LU-12) is ubiquitously expressed during mouse embryogenesis and in adult organs while human hematopoietic tissues showed differential expression. The mouse and human proteins are highly conserved containing two RNA recognition motifs (RRMs) and other domains associated with RNA binding and protein-protein interactions. A database search identified related proteins in human, rat, *C. elegans* and *S. pombe* including the 3p21.3 co-deleted gene, LUCA15. Recombinant proteins containing the RRMs of DEF-3(g16/NY-LU-12) and LUCA15 specifically bound poly(G) RNA homopolymers *in vitro*. These RRMs also show similarity to those of the Hu protein family. Since anti-Hu RRM domain antibodies are associated with an anti-tumor effect and paraneoplastic encephalomyelitis, we tested sera from Hu syndrome patients with the RRMs of DEF-3(g16/NY-LU-12) and LUCA15. These were non-reactive. Thus, DEF-3(g16/NY-LU-12) and LUCA15 represent members of a novel family of RNA binding proteins with similar expression patterns and *in vitro* RNA binding characteristics. They are co-deleted in some lung cancers and immunologically distinct from the Hu proteins.

Keywords: granulocyte differentiation; zinc finger; POZ domain

Introduction

Loss of heterozygosity involving 3p is one of the most frequent alterations in various epithelial neoplasms. Within this chromosomal arm, 3p21.3 appears to be an important target (Daly *et al.*, 1993; Wei *et al.*, 1996; van den Berg *et al.*, 1996) (and references therein). We previously identified a 3p21.3 homozygous deletion in the small-cell lung carcinoma (SCLC) cell line NIH-H1450 (Roche *et al.*, 1996) which overlapped two other deletions in the SCLCs, NIH-H740 and GLC20. In that report, we described the semaphorin, H.SemaIV, from the common deletion region and an unrelated cDNA, L1-204, from the site of the telomeric GLC20 breakpoint. Although L1-204 was evolutionarily conserved, it did not contain a significant open reading frame. However, GRAIL predicted one coding segment which corresponded to a 74 bp exon-trapping product. In an independent set of experiments using a retroviral gene-trap vector to identify differentially expressed genes during granulocytic and macrophage differentiation of murine FDCPmix-A4 cells, we isolated sequences homologous to the trapped exon. This gene, initially called def-3 (Acc No X96701), was found to be downregulated during granulocytic differentiation while persisting during macrophage development (Hotfilder *et al.*, data to be presented elsewhere).

Sequences from a partial cDNA, corresponding to the same gene, had been deposited as g16 (#U50839) nearly simultaneously with our initial submission. Recently, Gure *et al.* (1998) identified NY-LU-12/g16 by screening a cDNA expression library using an autologous antibody from a patient with adenocarcinoma of the lung. Of interest, these authors found that two of twenty-one lung cancer patients had reactive antibodies although no further characterization of the antibodies or biochemical analysis of the protein was performed.

In this report, we have examined the human and mouse DEF-3(g16/NY-LU-12) sequences and their RNA expression patterns. We also examined the expression of LUCA15, a highly similar and co-deleted gene located immediately adjacent to DEF-3(g16/NY-LU-12) in 3p21.3. DEF-3(g16/NY-LU-12) and LUCA15 encode predicted RNA binding proteins since each contains two RNA recognition motifs (RRMs). While similar to *Drosophila sex-lethal* and *elav*, among mammalian proteins the RRMs of DEF-3(g16/NY-LU-12) and LUCA15 are most like those of the human Hu protein family. Interestingly, RRMs from Hu proteins are antigenic and associated with paraneoplastic encephalomyelitis and anti-tumor activity in some patients with SCLC. Antibodies from Hu-syndrome patients are directed against the RRM domains and cross-react with multiple Hu protein family members (reviewed in Okano and Darnell, 1997). Therefore, we tested the cross-reactivity of DEF-3(g16/NY-LU-12) and LUCA15 RRMs with Hu antisera and found them to be immunologically distinct.

*Correspondence: HA Drabkin
Received 12 June 98; revised 29 October 1998; accepted 11 December 1998

As an initial characterization and confirmation that the predicted RRMs actually bind RNA, we performed *in vitro* studies with RNA homopolymers and recombinant RRM polypeptides from DEF-3(g16/NY-LU-12) and LUCA15. Both proteins specifically bound poly(G) RNA, in contrast to HuD which bound poly(A). With their overall similarity, this suggests they might have common targets *in vivo*. Finally, a database analysis indicated that DEF-3(g16/NY-LU-12) and LUCA15 are markedly similar to the rat RNA binding protein, S1-1, and its apparent human homologue (KIAA0122). Additional related proteins were identified in *C. elegans* and *S. pombe*. Thus, DEF-3(g16/NY-LU-12) and LUCA15 are members of an evolutionarily conserved family of RNA binding proteins.

Results

Cloning and mapping of DEF-3(g16/NY-LU-12)

Two approaches led us to the identification of DEF-3(g16/NY-LU-12). A mouse sequence of 450 bp (def-3, Acc No X96701) was isolated from the multipotent cell line, FDCPmix-A4, in a retroviral gene-trap approach to identify differentially regulated genes during granulopoeisis (Hotfilder *et al.*, in preparation). Independently, a positional cloning approach involving an apparent cluster of breakpoints near the telomeric boundary of an SCLC 3p21.3 homozygous deletion led to isolation of a trapped exon from an incompletely spliced cDNA (L1-204) (Roche *et al.*, 1996). This exon was nearly identical to sequences from the mouse clone which, in turn, identified the partial cDNA in GenBank, g16 (Acc No U50839). As initially deposited, g16 lacked the trapped exon from L1-204. Both the human and mouse full-length cDNAs were subsequently isolated (Materials and methods). A primer pair, H29465, from the 3′ end of DEF-3(g16NY-LU-12) was found to be commonly deleted in each of three SCLC cell lines, i.e., NCI-H740, GLC20 and NCI-H1450 (not shown). However, a Southern blot using a probe from 359–1037 bp of DEF-3(g16/NY-LU-12) showed a homozygous deletion only in NCI-H1450 (Figure 1). Thus, DEF-3(g16/NY-LU-12) is directly interrupted by both the NCI-H740 and GLC20 deletions.

Sequence analysis of DEF-3 identifies a novel protein family

The human cDNA sequence (Acc No AF069517) of DEF-3(g16/NY-LU-12) predicts a 1123 amino acid nuclear protein of 129 kDa. Comparison with the mouse sequence (Acc No AJ006486) (Figure 2a) shows they are 89% identical and 93% similar at the amino acid level. Database searches identified related proteins in human (LUCA15 and KIAA0122), rat (S1-1), *C. elegans* and *S. pombe*. All exhibit a similar domain structure including the order in which these domains occur, thus defining a novel protein family. Table 1 shows the similarities and amino acid positions for these various domains. Those domains common to each of the family members are the RRMs, zinc fingers and the C-ter region. The latter occupies approximately

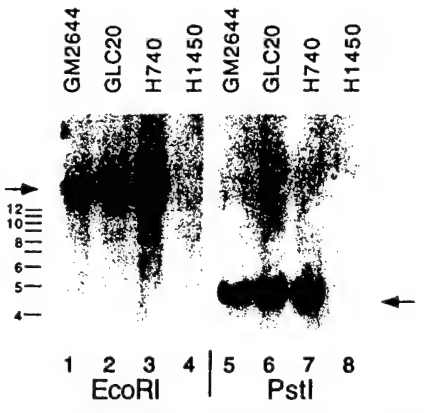


Figure 1 A portion of DEF-3(g16/NY-LU-12) is retained in cell lines defining the telomeric border of the 3p21.3 common deletion in SCLC cell lines. DNA samples from a normal human control cell line (GM2644) and three SCLC cell lines (GLC20, NCI-H740 and NCI-H1450) with known homozygous deletions were digested with *Eco*RI or *Pst*I. The resulting Southern blot was hybridized with a 678 bp DEF-3(g16/NY-LU-12) cDNA probe from bp 359–1037

70 amino acids and is rich in glycine, serine/threonine and arginine/lysine residues, a feature noted in a variety of RNA binding proteins (Birney *et al.*, 1993). A comparison of the various C-termini is shown in Figure 2b.

Two domains are unique to DEF-3(g16/NY-LU-12). The first is the decamer repeat, located from amino acids 71–330, which occurs more than 20 times. The most invariant amino acids in the repeats are DFRGRD with adjacent uncharged residues on either side. In certain repeats, RGR is substituted by RGG, a recognized motif (the RGG box) in some RNA binding proteins (Burd and Dreyfuss, 1994). Thus this region, like the C-terminus, may participate in RNA binding. The second unique domain involves amino acids 14–65 which shows 48% similarity/23% identity with the POZ (poxvirus and zinc finger) domain of ZF5 (Acc No D89859). This degree of similarity, while perhaps not striking, is comparable to that observed among other POZ domains. Moreover, in DEF-3(g16/NY-LU-12), the second zinc finger belongs to the C_2H_2 class ($CX_{2-4}CX_3FX_5LX_2HX_{3-4}H$). Zinc fingers of this class can be divided into subfamilies based on their N-terminal domains (reviewed in Grondin *et al.*, 1996). One of these subfamilies is defined by N-ter POZ domains, suggesting that this prediction is more likely to be correct. POZ domains are implicated in protein-protein interactions (Bardwell and Treisman, 1994; Dong *et al.*, 1996) and have been reported to inhibit binding of target sequences by zinc fingers (Kaplan and Calame, 1997; Numoto *et al.*, 1993).

Expression of DEF-3(g16/NY-LU-12) and LUCA15

Northern blots for DEF-3(g16/NY-LU-12) were performed with RNAs of both mouse and human derived tissues. A 450 bp mouse DEF-3(g16/NY-LU-12) probe corresponding to 1225–1674 bp of the mouse gene was used in Figure 3a. During mouse embryogenesis (part a), and in adult non-hematopoietic

tissues (part b), expression appears constitutive. However, using the same probe against human hematopoietic RNAs, DEF-3(g16/NY-LU-12) expression is highest in certain lymphoid tissues (lymph node and thymus) while markedly lower in bone marrow. These results are consistent with our initial findings which suggested that DEF-3(g16/NY-LU-12) expression was reduced during granulocytic differentiation. In peripheral blood leukocytes and also in fetal liver, two bands were observed. In support of this, several

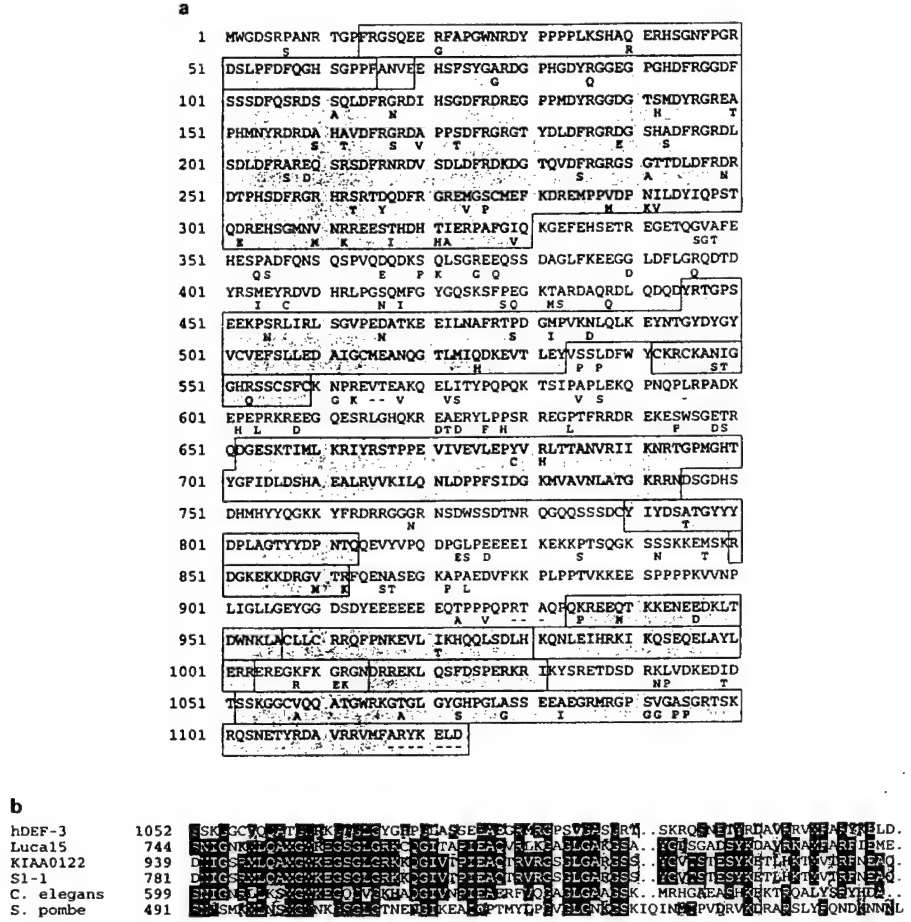


Figure 2 (**a**) Amino acid sequence of the human and mouse DEF-3(g16/NY-LU-12) genes. The human sequence is given in the upper line and differences in the mouse sequence are noted below. The shaded boxes demarcate the amino acid positions (given in parentheses) of following consecutive domains: POZ (14-65), decamer repeats (70–330), RRMI (445–533), ZnFl (542–559), RRM2 (652–744), octamer repeats (790–813), KEKE motif (850–862), coiled-coil (934–956), ZnF2 (957–980), coiled-coil (981–1003), nuclear localization signal (1015–1031) and the Gly, Ser/Thr, Arg/Lys rich C-ter (1052–1123). (**b**) Alignment of C-termini from DEF-3(g16/NY-LU-12) and family members. Genbank accession numbers are LUCA15 (#U23946), KIAA0122 (#D50912), S1-1 (#D83948), *C. elegans* gene (1947021 in #AF000263) and *S. pombe* gene (2104448 in #Z95396)

Table 1 The various domains are listed across the top

	POZ	*Decamer repeats*	*RRM1*	*ZnF1*	*RRM2*	*Octamer repeat*	*KEKE motif*	*Amphipathic helix*	*Coiled coil*	*ZnF2*	*Coiled coil*	*Nuclear local*	*Gly-Rich*
hDEF-3	14–65	70–330	445–533	542–559	652–744	790–813	850–862		934–956	957–980	981–1003	1015–1031	1052–1123
			58%	38%	61%	67%	77%		57%	93%	80%	63%	61%
Luca15			112–177	187–204	231–315	456–498	523–535	536–553	614–648	649–671	672–711	707–724	744–815
			100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
KIA-			225–289	300–317	375–464	646–688	714–726	727–744		837–863	864–905	900–917	940–1010
A0122			75%	83%	82%	86%	85%	100%		93%	87%	100%	79%
S1-1			61–131	142–159	217–307	488–530	556–568	569–586		680–706	707–747	735–759	782–852
			74%	83%	81%	86%	85%	100%		96%	87%	100%	79%
C. Ele-			5–50	66–83		299–340		375–392		484–509	52–537		599–669
gans			65%	72%		55%		78%		65%	47%		67%
S. Pombe			69–151	161–178						403–428	370–394		491–565
			53%	56%						56%	57%		52%

The numbers in the columns refer to the amino acid positions of the domains in their respective proteins. The per cent values indicate the similarity of each domain compared to LUCA15, initially identified as Gene15 (Wei *et al*., 1996), which was used as the base of comparison since it has an overall higher similarity to the others and because it contains an amphipathic domain present in some members, but absent in DEF-3(g16/NY-LU-12). The POZ and decamer repeats are unique to DEF-3(g16/NY-LU-12)

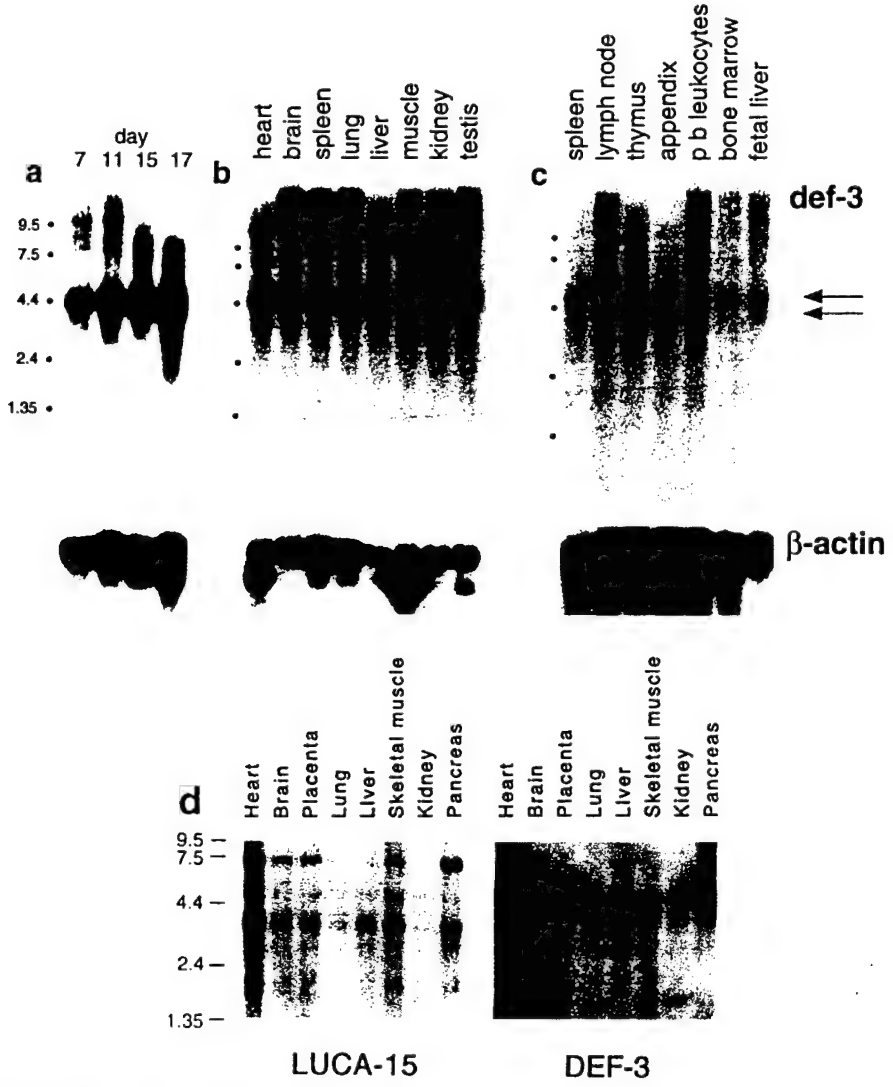


Figure 3 DEF-3(g16/NY-LU-12) expression in mouse and human tissues. RNA was isolated from fetal mice at the indicated gestational times (**a**) or from adult tissues (**b**). Human RNAs from the indicated tissues are shown in (**c**). The filters were hybridized with a mouse DEF-3(g16/NY-LU-12) probe corresponding to bp 1225–1674. The arrows indicate two RNA products present in human samples. β-actin was used as control (bottom panels). A separate Clontech blot (**d**) was sequentially hybridized with human LUCA15 and DEF-3(g16/NY-LU-12) probes corresponding to bp 2055–2561 and 1037–1015, respectively (see Materials and methods)

splicing variants were identified in the cDNA database (dbEST) which significantly alter the predicted protein (e.g., EST AA934109, AA442117).

In additional experiments, DEF-3(g16/NY-LU-12) expression was directly compared to LUCA15 by sequential hybridizations of the same filter (Clontech). Overall, the expression patterns for both DEF-3(g16/NY-LU-12) and LUCA15 are similar. Figure 3d shows that both genes are most highly expressed in heart while expression in lung and kidney was barely detectable. This near absence of expression in lung and kidney was partially due to less RNA present in those lanes as determined by a control G3PDH probe (not shown). For DEF-3(g16/NY-LU-12), the relative intensities of the ~4 and 5 kb bands vary considerably among different tissues. Two bands are also evident for LUCA15, although their nearly constant relative intensities and considerable size difference suggests that the larger band is incompletely spliced. To further extend the expression analysis, a dot blot (Clontech) representing 50 tissues was sequentially hybridized and quantitated by phosphorimaging. For comparison, the counts from all 50 tissues were averaged and each compared to the mean. Like the Northern analysis, expression levels of the two genes were quite similar (not shown). Highest levels were seen in adult thymus (DEF-3(g16/NY-LU-12), 224% of the mean; LUCA15, 179%) and fetal kidney (DEF-3(g16/NY-LU-12), 183%; LUCA15, 176%). In contrast, levels in the fetal thymus and adult kidney counterparts were among the lowest expressing tissues: fetal thymus (DEF-3(g16/NY-LU-12), 72%; LUCA15, 79%) and adult kidney (DEF-3(g16/NY-LU-12), 67%; LUCA15, 79%). These results suggest that differential expression of DEF-3(g16/NY-LU-12) and LUCA15 occurs during development of non-hematopoietic organs as well. The most marked difference between the two genes was in adult frontal lobe where DEF-

3(g16/NY-LU-12) was 135% of the mean compared to 55% for LUCA15.

Binding of RRM domains to RNA and reactivity with anti-Hu sera

Sequence similarities in the RRM domains of RNA binding proteins have been reported to predict functional similarities among these proteins (Kim and Baker, 1993). BLAST searches and tertiary structural alignments showed the DEF-3(g16/NY-LU-12) and LUCA15 RRMs to be most closely related to RRMs from *Drosophila sex-lethal (Sxl)*. Among mammalian proteins, the most related RRMs are those from the Hu protein family. Hu proteins bind the 3′ UTRs of some growth factor and cytokine messages resulting, where known, in increased stability (Gao *et al*., 1994; Levine *et al*., 1993). In patients with SCLC, anti-Hu antibodies (which cross-react with multiple Hu protein family members) appear to be associated with an anti-tumor activity and in some cases result in paraneoplastic encephalomyelitis (Okano and Darnell, 1997; Graus *et al*., 1997). Figure 4 shows a comparison of the second RRM from DEF-3(g16/NY-LU-12) and LUCA15 to the first RRM of HuC. The overall structural organization, $\beta\alpha\beta\beta\alpha\beta$, is conserved (Burd and Dreyfuss, 1994). The β sheets containing the conserved octamer (RNP1) and hexamer (RNP2) motifs are believed to provide a general RNA binding surface whereas more specificity is provided by the variable loop segments and boundaries of the RRM domain (Burd and Dreyfuss, 1994).

Recombinant proteins containing the RRMs of DEF-3(g16/NY-LU-12) and LUCA15 were produced in bacteria (Materials and methods). After affinity chromatography, the proteins appeared either essentially pure by Coomassie staining, or they contained one or more faint lower bands which reacted with anti-T7 Tag antibodies and therefore represented truncated products (i.e., the T7 tag epitope is encoded by the pRSET vector in the N-terminus of the fusion protein). The predicted sizes of the recombinant proteins were ~44, 33 and 45 kDa for DEF-3(g16/NY-LU-12), LUCA15 and HuD, respectively, which matched their migration in SDS-polyacrylamide gels. The DNA constructs were used to produce *in vitro* transcribed/translated ^{35}S-labelled proteins for binding to RNA homopolymers and the purified proteins from bacteria were used to test cross-reactivity with anti-Hu sera. As a negative control for RNA binding, we prepared labelled firefly luciferase protein (61 kDa) using the template provided in the TnT kit (Materials and methods).

The results of the RNA binding studies are shown in Figure 5. Preferential binding was observed by the DEF-3(g16/NY-LU-12) and LUCA15 RRM domains to poly(G) RNA (lanes 14, 16 and 20, respectively) at near physiologic salt concentration (0.1 M NaCl, pH 7.5). Of note, the DEF-3(g16/NY-LU-12)/LUCA15 family member SI-1 (rat homologue of human KIA0122) was previously shown to bind G and U polyribonucleotides (Inoue *et al*., 1996). In contrast, HuD was found to bind poly(A) RNA (lane 6) in agreement with its reported binding to specific A + U-rich substrates (Okano and Darnell, 1997; Gao *et al*., 1994). No binding was observed by the 61 kDa firefly luciferase (lanes 27–33). For DEF-3(g16/NY-LU-12), we tested its binding to poly(G) RNA at high salt (0.5 M NaCl, pH 7.5) and found it was inhibited (lane 17). Finally, to exclude the possibility that the observed binding was due to a contaminating intermediate protein from the reticulocyte lysate, we performed a UV cross-linking experiment. In this procedure, only proteins in direct contact with nucleic acid are covalently cross-linked to it by UV exposure. After incubation of DEF-3(g16/NY-LU-12) with poly(G) RNA, UV exposure resulted in significantly less recoverable protein from the beads after boiling in SDS-containing sample buffer (lane 19). Based on the presence of the structurally conserved RRM domains and these biochemical results, we conclude that DEF-3(g16/NY-LU-12) is an RNA binding protein. No significant binding was observed with either single-strand or double-strand DNA (not shown).

Because the RRMs of DEF-3(g16/NY-LU-12) and LUCA15 show similarity to those of the Hu proteins, we asked whether anti-Hu antibodies would be cross-reactive. High titer sera from ten patients with the Hu syndrome were tested by Western blotting against the recombinant RRMs from DEF-3(g16/NY-LU-12), LUCA15 and, as a positive control, HuD. Four representative samples for DEF-3(g16/NY-LU-12) are shown in Figure 6. While all the anti-Hu sera were reactive to HuD, no cross-reaction was detected against DEF-3(g16/NY-LU-12). Identical results were obtained for LUCA15 (not shown). Control sera showed no reactivity against any of the proteins (not shown). Thus, while the RRMs of DEF-3(g16/NY-LU-12) and LUCA15 are similar to those from HuD, they are immunologically distinct.

Discussion

RNA-binding proteins regulate basic aspects of RNA metabolism such as capping, polyadenylation, transport, splicing, stability and translation (reviewed in Burd and Dreyfuss, 1994; Singh *et al*., 1995). The presence of two RNA recognition motifs in DEF-3(g16/NY-LU-12) suggested it would bind RNA.

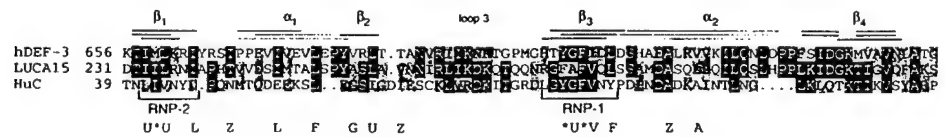


Figure 4 Alignment of RRM domains from DEF-3(g16/NY-LU-12), LUCA15 and HuC. Black and shaded boxes indicate identity and similarity, respectively. Lines above the alignment indicate the predicted secondary structures. The topmost line corresponds to DEF-3(g16/NY-LU-12) and so-forth. Conserved residues as given by Birney *et al*. (1993) are listed below the alignment where U indicates uncharged and Z indicates uncharged, S, or T

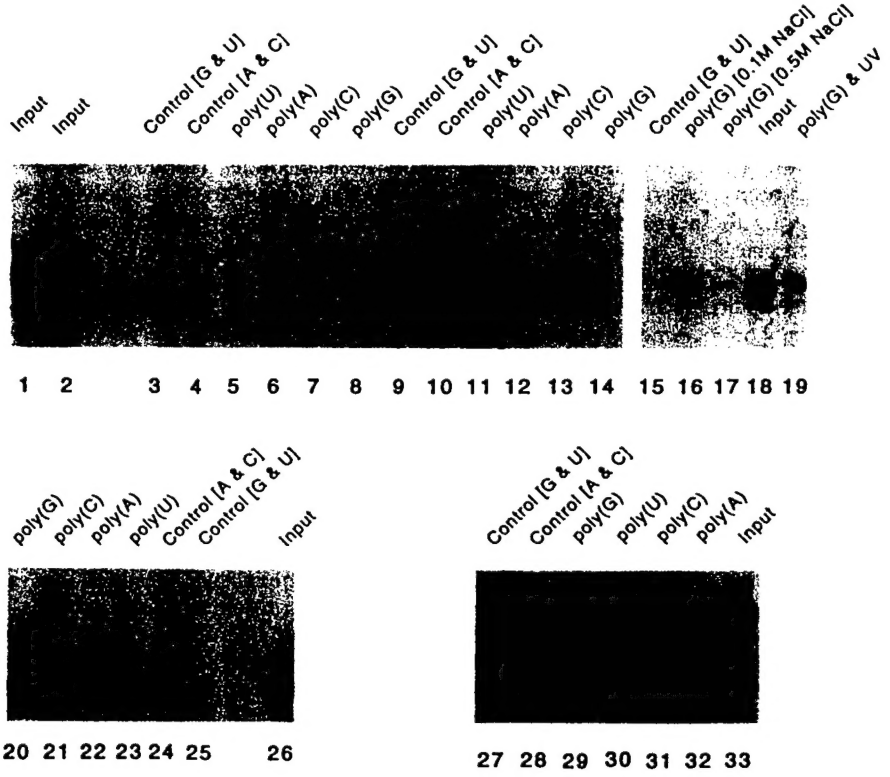


Figure 5 Results of ^{35}S-labelled protein binding to RNA homopolymers. Lanes 1 and 2 contain approximately 10% of the TnT reactions from HuD and DEF-3(g16/NY-LU-12), respectively. Lanes 3–8 represent HuD binding to the indicated RNA homopolymers or corresponding matrices (see Materials and methods). The DEF-3(g16/NY-LU-12) results are in lanes 9–14 and 15–19. As can be seen, HuD binds preferentially to poly(A) RNA (lane 6). DEF-3(g16/NY-LU-12) binds poly(G) RNA (lanes 14 and 16) which is inhibited by high salt (lane 17). The results of UV cross-linking are shown in lane 19. The LUCA15 results are shown in lanes 20–26 with preferential binding to poly(G). The negative control reactions with firefly luciferase are in lanes 27–33

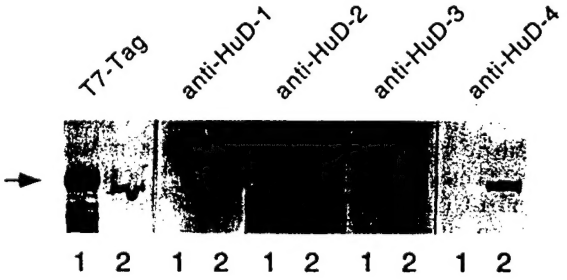


Figure 6 Western blots showing the reactivity of sera from four representative patients with the Hu syndrome to recombinant HuD protein in lane 2. In contrast, no reaction is seen against recombinant DEF-3(g16/NY-LU-12) protein in lane 1. Both DEF-3(g16/NY-LU-12) and HuD proteins contain the T7 epitope tag as shown in the first pair of lanes. Considerable more of the DEF-3(g16/NY-LU-12) protein was loaded in order to enhance detection of cross-reacting antibodies. Identical results were observed with LUCA15 (not shown)

RRMs comprise 80–90 amino acids and include an octapeptide, RNP1 and a hexapeptide, RNP2 located in the first and third β-sheets (Singh *et al*., 1995; Kenan *et al*., 1991; Kim and Baker, 1993; Birney *et al*., 1993). From X-ray crystallographic and other studies, each RNP binds 5–7 specific nucleotides (Oubridge *et al*., 1994). The length of intervening loops and the distribution of basic amino acids adds to the stability of RNA-binding. Using recombinant proteins produced by *in vitro* transcription/translation, we have demonstrated that the DEF-3(g16/NY-LU-12) RRMs, as well as LUCA15, are capable of specifically binding poly(G) RNA. For DEF-3(g16/NY-LU-12), this binding occurred at physiologic but not elevated salt concentrations and the UV cross-linking studies confirmed that the binding was directly due to DEF-3(g16/NY-LU-12). In similar studies, the related rat S1-1 protein was shown to bind poly(G) and poly(U) RNA (Inoue *et al*., 1996). While the homopolymer studies demonstrate that these proteins can bind RNA they are not designed to identify more specific *in vivo* targets.

Interestingly, similarity among RRMs is associated with functional similarity (Kim and Baker, 1993). For instance, the RRMs of proteins associated with hnRNAs are more similar to each other than to the RRMs of proteins involved in differentiation (Kim and Baker, 1993). In vertebrates, Hu-family proteins bind to A+U rich destabilizing elements in the 3'UTRs of certain growth factor and cytokine messages affecting message stability (Gao *et al*., 1994; Levine *et al*., 1993). Thus for example, we might expect DEF-3(g16/NY-LU-12) to be involved in the regulation of splicing or message stability instead of RNA transport. Consistent with this possibility, we found DEF-3(g16/NY-LU-12) was down-regulated during granulocytic differentiation in FDCPmix-A4 cells while it persisted during macrophage formation (Hotfilder *et al*., in preparation). During hematopoiesis, certain key transcription factors are regulated by

alternative splicing. For example, Ikaros is an alternatively spliced zinc finger transcription factor required for lymphoid development which, when knocked out, results in an excess of monocytes and macrophages (Wu *et al.*, 1997). In our Northern analysis, DEF-3(g16/NY-LU-12) expression was considerably higher in certain lymphoid tissues (thymus and lymph node) compared to bone marrow which is predominantly composed of myeloid and erythroid lineages. DEF-3(g16/NY-LU-12) is predicted to be a nuclear factor based on its nuclear localization signal. Also, its coiled-coil, zinc finger, KEKE and putative POZ domains suggests it might interact with several proteins.

DEF-3(g16/NY-LU-12) and the related LUCA15 are co-deleted in three SCLC cell lines with homozygous deletions involving 3p21.3. We have shown that DEF-3(g16/NY-LU-12) is located at the telomeric boundary of the common deletion region defined by these SCLCs. The Hu syndrome is a paraneoplastic neuropathy and encephalomyelitis associated with high-titre antibodies directed against the RRMs of Hu-proteins (Dalmau *et al.*, 1992). In the majority of reported cases, the associated tumor was lung cancer and most were small-cell carcinomas, although adenocarcinoma was also observed (Dalmau *et al.*, 1992). High titre anti-Hu antibodies appear to confer an anti-tumor effect (Dalmau *et al.*, 1992) and even low-titre antibodies, which occur in 16% of SCLC patients, have apparent anti-tumor activity (Graus *et al.*, 1997). It is interesting in this regard that while no mutations in HuD were observed in lung cancers (Sekido *et al.*, 1994), another Hu family gene, Hel-N1, is located in a novel deletion region on 9p (Cairns *et al.*, 1997). Therefore, we were interested in determining if anti-Hu antibodies cross-reacted with the DEF-3(g16/NY-LU-12) or LUCA15 RRMs. Our findings demonstrate they do not.

During preparation of this manuscript, Güre *et al.* (1998) identified NY-LU-12/g16, the same gene as DEF-3(g16/NY-LU-12), on the basis of an autologous antibody from a patient with adenocarcinoma of the lung. No tumor specific mutation was identified in NY-LU-12/g16 as a possible explanation for the reactive antibodies although LUCA15 was not examined. Moreover, these authors found that two of 21 lung cancer patients had anti-NY-LU-12/g16 antibodies. The cDNA clone which reacted with the patient's serum was reported to contain 952 nucleotides of C-ter coding sequence. While this would exclude the RRM domain as targets, the C-ter region demonstrates considerable similarity among DEF-3(g16/NY-LU-12) related family members. Thus, it is possible that anti-DEF-3(g16/NY-LU-12) antibodies might be cross-reactive against LUCA15. We previously reported that two of 19 adenocarcinomas contained homozygous deletions using a FISH probe from the H.SemaIV locus, also within the 3p21.3 deletion (Varella-Garcia *et al.*, 1997). If anti-DEF-3(g16/NY-LU-12) or anti-LUCA15 antibodies have an anti-tumor effect, then a co-deletional event would remove both potential cross-reactive antigens. Also, the deletions we observed occurred in only a portion of the tumor cells and would likely be missed by PCR. Thus, it is possible that an anti-tumor effect from antibodies, or other forms of immunity, directed against DEF-3(g16/NY-LU-12) or LUCA15 might occur in some patients with lung cancer.

Materials and methods

Cloning of DEF-3(g16/NY-LU-12)

A 74 bp trapped exon from the cDNA clone L1-204 (Roche *et al.*, 1996) was 99% identical to a partial mouse def-3 sequence (Acc No X96701) which, in turn, provided linkage to the human g16 partial cDNA deposited by Latif *et al.* (U50839). From analysis of dbEST, clone zq05f07.s1 from a Stratagene muscle cDNA library was obtained and completely sequenced. For the 5′ end, nested gene specific primers 5′-TCC CCT CCT CGA TAG TCA CC and 5′-CCA TAG CTG AAA GAA TGC TCC were used with T3 and T7 vector primers (each at 1 μM) to amplify products from DNAs prepared from a fetal brain cDNA library (Statagene, #936206). PCR reactions were performed under 65°C-55°C touch-down conditions (ΔT of -1.0°C per 2 cycles for 20 cycles, followed by 15 cycles at 55°C). Products were cloned into a T-vector prepared from pBluescript II KS and sequenced on an ABI377 through the University of Colorado Cancer Center DNA Sequencing Core.

The mouse DEF-3(g16/NY-LU-12) sequence from nucleotides 1225-3593 (Acc No AJ006486) was assembled from our original clone, mdef-3 (Acc No X96701), and two ESTs (Acc Nos AA268319 and AA086694); the latter was obtained from the UK HGMP Resource Centre (Cambridge) and its complete sequence determined. A gap between the ESTs was bridged by RT–PCR using brain mRNA isolated from strain C57/Bl/6 and gene specific primers GSP3 (5′-ATG AAC TCA GGA AAC GGC ACG) and GSP6 (5′-AGG CTT CCC CTC ACT GGT ACT TT). The products were diluted 1 : 10 and reamplified under the same conditions using internal primers GSP4 sense, 5′-AGG GAA GGG CT-T ACT TTC CGA and GSP5 antisense, 5′-TTG CTA TTT -GAC TTT CCT TGA CTG GTA. For these experiments, RNA was prepared using the QuickPrep Micro mRNA Purification Kit (Pharmacia, Cat No 27-9255-01) and reverse transcription was performed on random primed template as per the 5′ RACE system Version 2.0 protocol (Gibco, Cat No 18374-058). PCR conditions included 30 cycles for each round, an annealing temperature of 59°C and Exo$^-$ Pfu DNA polymerase and buffer (Stratagene, Cat No 600163). The final products were cloned into the T-vector, pCR2.1 (Invitrogen, Cat No K2030-01) and sequenced using the LI-COR 4200 System (MWG-Biotech, UK).

Nucleotides 1–264 were obtained by RACE using nested gene specific primers GSP25 [5′-TCC ATA GCT GAA AGA ATG CTC C-3′] and GSP32 [5′-AAA GGG AGG TCC TGA ATG CCC CTG GAA ATC-3′] corresponding to nt 411–432 and nt 370–399 respectively, of the human gene. The forward RACE primer was the Abridged Anchor Primer (Gibco, 5′-GGC CAC GCG TCG ACT AGT ACG GGI IGG GII GGG IIG-3′). A separate Race primer was used for the second stage PCR; the Abridged Universal Anchor Primer (5′-GGC CAC GCG TCG ACT AGT AC-3′). Nucleotides 654–1224 were similarly obtained by RACE using gene specific primers GSP1 (5′-TCT GGA ACC CCA CTT AAT CGA ATA A-3′) and GSP2 (5′-TTC TTT AAA CAG ACC AGC GTC TGA A-3′). For both experiments, PCR conditions consisted of [94°C 2 min, (94°C × 1 min, 55°C 1 min, 68°C × 2 min) for 30 cycles, 72°C × 10 min] with the LA Taq and buffer System Takara Shuzo Co., Biowhittaker UK Ltd). Final products were cloned into the pGEMT vector for sequencing. The remaining nucleotides 265–653 were from mouse EST sequence (Acc No AF064939) kindly provided by Dr C Kappen (Scottsdale, AZ, USA).

Cloning and purification of bacterially produced RRM proteins

The following PCR primers, forward (5'-GGA ATT C-GA TGC CCA ACG GGA C) and reverse (5'-AAC-TAA AGC TTA GCA TAA TCT GGA ACA TCA TAT-GGA TAG TGG TCA GAA TGG TC) containing *Eco*RI and *Hind*III sites, respectively, were used to amplify a segment containing both RRMs from amino acids 435–752 of DEF-3(g16/NY-LU-12). The reverse primer also contained an HA-tag.DNA from a colony purified cDNA clone of zq05f07.s1 served as template. The PCR products were digested with *Eco*RI and *Hind*III, cloned in-frame using the bacterial expression vector, pRSETB (Invitrogen), and verified by DNA sequencing. This construct which includes vector derived T7 Tag and polyhistidine segments along with the HA-tag and 17 amino acids of downstream vector sequences is expected to produce a protein of approximately 44 kDa. Recombinant protein was produced in BL21lys S cells (Novagen) and purified using a Talon resin (Clontech) with imidazole-based buffers for native 6xHis protein purification as recommended by the manufacturer (Protocol #PT1320-1).

Similarly, the entire HuD coding sequence and the LUCA15 sequence from amino acids 90–357 which includes both RRM domains were PCR amplified from fetal brain cDNA library DNA as above and cloned into pRSETB. Sequence verified constructs were expressed and purified as above except that recombinant HuD protein appeared to require purification in denaturing conditions for optimal yields. The primer sequences were HuD_for (CGG GAT CCG GTT ATG ATA ATT AGC ACC ATG-G), HuD_rev (CTC CCG CTC GAG TCA GGA CTT-GTG GGC TTG G), LUCA15_AA90_for (CGG GAT CC-G ATC AGT GAC GAG AGG GAG AGC) and LUCA 15_AA357_rev (CTC CCG CTC GAG TCA GTC AAC-ACT GAC TCC TTC ACC AC). The predicted sizes of the recombinant proteins were 45 kDa for HuD and LUCA15, 33 kDa.

Western blots and testing of anti-HuD antisera

Recombinant DEF-3(g16/NY-LU-12), LUCA15 and HuD proteins were separated on 12% SDS-polyacrylamide (29:1) gels and transferred to nitrocellulose filters. Blocking was performed with 1% (wt/vol) I-BLOCK (Tropix, Bedford, MA, USA) overnight, washed three times (5 min each) in phosphate buffered saline (PBS), and incubations with primary antibodies were carried out in PBS containing 0.1% (vol/vol) Tween-20 for 1 h at room temperature. The filters were washed as above, then similarly incubated with alkaline phosphatase conjugated secondary antibodies. After final washing the filters were developed using Pierce (Rockford, IL, USA) 1-step NBT/BCIP. The primary antibodies were: (1) T7 Tag (Novagen, Madison, WI, USA, #69522-3) at 1:10 000 dilution, a mouse monoclonal directed at the T7 gene 10 leader sequence encoded by pRSEB; (2) 12CA5 (Boehringer-Mannheim, #1583816) at 1 μg/ml, a mouse monoclonal directed at the HA peptide; (3) human sera from ten SCLC patients with the Hu syndrome, previously documented to contain anti-HuD antibodies, used at 1:1000 dilution. The secondary antibodies were: (1) rabbit anti-mouse AP-F(ab')$_2$, used at 1:5000 dilution (Zymed, San Francisco, CA, USA, #61-6322); (2) goat anti-human IgG used at 1:2500 dilution (Jackson ImmunoResearch Labs, #109-055-003).

Binding of RRM proteins to RNA Homopolymers

The pRSETB constructs described above were used as templates for the production of ^{35}S-labelled protein by *in vitro* transcription/translation with the TnT System (Promega, Madison, WI, USA) under conditions specified by the manufacturer. As a negative control for RNA binding, firefly luciferase (61 kDa) was translated using the template provided in the TnT kit. Binding of ^{35}S-labelled proteins to poly(A) (Sigma #P-8708), poly(U) (Sigma #P-8563), poly(G) (Sigma #P-1908) and poly(C) (Sigma #P-9827) RNA homopolymers was performed as utilized by Ohno *et al.* (1994). Binding was similarly examined using single-stranded (Sigma #D-8273) and double-stranded (Sigma #D-8151) DNA resins. The control resins were Sepharose CL-4b-200 (Sigma), Cellex 410 cellulose powder (BioRad Labs) and polyacrylhydrazido-agarose (Sigma #P-9282). Ten per cent of the *in vitro* transcription/translation reaction in 0.5 ml of 0.1 M NaCl/10 mM Tris (pH 7.5) was added to 25 μl of each resin previously equilibrated in 0.1 M NaCl/10 mM Tris (pH 7.5) and another 10% was saved for direct loading on the gel. The mixture was rotated at 4°C for 1 h, then briefly pelleted and the supernatant removed. The resin was briefly washed two times with 0.5 ml of 0.1 M NaCl/10 mM Tris (pH 7.5), then two more times for 20 min each. In a separate experiment with DEF-3(g16/NY-LU-12), binding to the poly(G) homopolymer was tested in the presence of 0.5 M NaCl/10 mM Tris (pH 7.5), but washed as above. Forty μl of 1×SDS–PAGE sample buffer was added to the washed resins and heated in a 100°C dry block for 5 min. Samples were electrophoresed in a 9% polyacrylamide gel, fixed in 10 volumes of 10% (vol/vol) glacial acetic acid for 20 min and neutralized in a dilute (4 mM) NaOH solution (until the bromophenol blue dye began to return to a blue color). The gel was soaked in 1 M salicylate for 30 min, dried and exposed to Kodak X-OMAT AR film. For the UV cross-linking experiments, following the binding and washing of ^{35}S-labelled DEF-3(g16/NY-LU-12) with the poly(G) homopolymer or control polyacrylhydrazido-agarose, the washed resin was placed in a petri dish and exposed to emissions from a 15-watt germicidal lamp (Sylvania G15T8, GTE Sylvania Inc, Danvers, MA, USA) at a distance of 4.5 cm for 10 min. Non-covalently bound DEF-3(g16/NY-LU-12) was recovered and analysed as above.

Cell lines, probes and hybridizations

The SCLC cell line GLC20 (Kok *et al.*, 1994) was kindly provided by Dr Charles Buys. The SCLC lines NCI-H740 and NCI-H1450 were obtained through the University of Colorado Lung Cancer SPORE and Drs Adi Gazdar and John Minna. DNA from the human cell line GM2644 was used as a normal control. The cell lines were grown in RPMI-1640 containing 10% fetal calf serum. For Southern blot analysis of DEF-3(g16/NY-LU-12), a 677 bp probe was generated using the primers 5'-CAC TTC CCT TTG ATT TCC AG (bp 359–378) and 5'-TAC AAG ATC CCA TCT CTC TGC C (bp 1037–1015) with the PCR conditions described above. For LUCA15, a PCR derived probe was obtained from bp 2055–2561 (accession #U23946) using the forward and reverse primers 5'-GAC CAC TTC TTG TGC TCC TTG and 5'-AGT CAG GCA CCA GCA ACA CT. The H29465 primers from the 3' portion of DEF-3(g16/NY-LU-12) (bp 2791–2871) were 5'-TTT CAG GAA AAT GC-C AGT GA and 5'-TTC CTT CTT CAC AGT AG-G AGG C. Southern blot hybridizations were performed at 65°C using a charged nylon membrane (Amersham) in Amasino buffer (Amasino, 1986). Washings were performed two times for 15 min at 65°C in 0.1×SSC/0.1% SDS. Northern blots containing approximately 2 μg of poly(A)$^+$ RNA per lane from human and mouse tissues, as well as a human RNA dot blot (RNA Master blot), were obtained from Clontech and utilized under conditions suggested by the supplier.

DNA sequence analysis

Protein motifs were found using the PRINTS protein motif fingerprint database (http://www.biochem.ucl.ac.uk/bsm/dbbrowser/PRINTS/PRINTS.html), Statistical Analysis of Protein Sequences at the Swiss Institute for Experimental Cancer Research (ISREC) (http://ulrec3.unil.ch/software/SAPS_form.html), and through alignments of DEF-3(g16/NY-LU-12) family members using the FASTA and BESTFIT programs from the GCG package. Coiled-coils were located using COILS at ISREC (http://ulrec3.unil.ch/software/COILS_form.html). Secondary structure was predicted using the Baylor College of Medicine SSP program (http://dot/imgen.bcm.tmc.edu:9331/psspredction/pssp.html), and tertiary structure alignments were done using the UCLA-DOE fold recognition server (http://fold/doe-mci.ucla.edu/Home). Database searches for homologs were conducted using the BLAST server at NIH.

Acknowledgements

We thank Dr C Korch for assistance with nucleotide sequencing through the auspices of the University of Colorado Cancer Center DNA Sequencing Core, supported by P30 CA 46934 and Dr C Kappen, Scottsdale, AZ for providing EST clone pCHR15-5. This investigation was supported by grants CA58187 and DAMD17-94J-4391 from the NIH and Department of Defense and the Special Trustees of the Middlesex Hospital, UCH and UCL Medical School (CRDC) and The Wellcome Trust.

References

Amasino R. (1986). *Analytic. Biochem.*, **152**, 304–307.

Bardwell VJ and Treisman R. (1994). *Genes Dev.*, **8**, 1664–1677.

Birney E, Kumar S and Krainer AR. (1993). *Nucl. Acids Res.*, **21**, 5803–5816.

Burd CG and Dreyfuss G. (1994). *Sci.*, **265**, 615–620.

Cairns P, Okami K, King P, Bonacum J, Ahrendt S, Wu L, Mao L, Jen J and Sidransky D. (1997). *Cancer Res.*, **57**, 5356–5359.

Dalmau J, Graus F, Rosenblum MK and Posner JB. (1992). *Medicine (Balt.)*, **71**, 59–72.

Daly MC, Xiang RH, Buchhagen D, Hensel CH, Garcia DK, Killary AM, Minna JD and Naylor SL. (1993). *Oncogene*, **8**, 1721–1729.

Dong S, Zhu J, Reid A, Strutt P, Guidez F, Zhong HJ, Wang ZY, Licht J, Waxman S and Chomienne C. (1996). *Proc. Natl. Acad. Sci. USA*, **93**, 3624–3629.

Gao FB, Carson CC, Levine T and Keene JD. (1994). *Proc. Natl. Acad. Sci. USA*, **91**, 11207–11211.

Graus F, Dalmau J, Rene R, Tora M, Malats N, Verschuuren JJ, Cardenal F, Vinolas N, Garcia del Muro J, Vadell C, Mason WP, Rosell R, Posner JB and Real FX. (1997). *J. Clin. Oncol.*, **15**, 2866–2872.

Grondin B, Bazinet M and Aubry M. (1996). *J. Biol. Chem.*, **271**, 15458–15467.

Gure AO, Altorki NK, Stockert E, Scanlan MJ, Old LJ and Chen Y-T. (1998). *Cancer Res.*, **58**, 1034–1041.

Inoue A, Takahashi KP, Kimura M, Watanabe T and Morisawa S. (1996). *Nucl. Acids Res.*, **24**, 2990–2997.

Kaplan J and Calame K. (1997). *Nucl. Acids Res.*, **25**, 1108–1116.

Kenan DJ, Query CC and Keene JD. (1991). *Trends Biochem. Sci.*, **16**, 214–220.

Kim YJ and Baker BS. (1993). *Mol. Cell Biol.*, **13**, 174–183.

Kok K, van den Berg A, Veldhuis PM, van der Veen AY, Franke M, Schoenmakers EF, Hulsbeek MM, van der Hout AH, de Leij L, van de Ven W and Buys CHCM. (1994). *Cancer Res.*, **54**, 4183–4187.

Levine TD, Gao F, King PH, Andrews LG and Keene JD. (1993). *Mol. Cell. Biol.*, **13**, 3494–3504.

Numoto M, Niwa O, Kaplan J, Wong KK, Merrell K, Kamiya K, Yanagihara K and Calame K. (1993). *Nucl. Acids Res.*, **21**, 3767–3775.

Ohno T, Ouchida M, Lee L, Gatalica Z, Rao VN and Reddy ES. (1994). *Oncogene*, **9**, 3087–3097.

Okano HJ and Darnell RB. (1997). *J. Neurosci.*, **17**, 3024–3037.

Oubridge C, Ito N, Evans PR, Teo CH and Nagai K. (1994). *Nature*, **372**, 432–438.

Roche J, Boldog F, Robinson M, Robinson L, Varella-Garcia M, Swanton M, Waggoner B, Fishel R, Franklin W, Gemmill R and Drabkin H. (1996). *Oncogene*, **12**, 1289–1297.

Sekido Y, Bader SA, Carbone DP, Johnson BE and Minna JD. (1994). *Cancer Res.*, **54**, 4988–4992.

Singh R, Valcarcel J and Green MR. (1995). *Sci.*, **268**, 1173–1176.

van den Berg A, Hulsbeek MF, de Jong D, Kok K, Veldhuis PM, Roche J and Buys CH. (1996). *Genes Chrom. Cancer*, **15**, 64–72.

Varella-Garcia M, Rao K, Rabenhorst SH, Wiest J, Drabkin H, Gemmill R, Anderson MW and Franklin WA. (1997). *Lung Cancer*, **18**, 140.

Wei MH, Latif F, Bader S, Kashuba V, Chen JY, Duh FM, Sekido Y, Lee CC, Geil L, Kuzmin I, Zabarovsky E, Klein G, Zbar B, Minna J and Lerman MI. (1996). *Cancer Res.*, **56**, 1487–1492.

Wu L, Nichogiannopoulou A, Shortman K and Georgopoulos K. (1997). *Immunity*, **7**, 483–492.